

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:39:30 ; Search time 1.38425 Seconds
(without alignments)
599.261 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPPMGSLKRRPQGECP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	74.1	435	1 TNR3_HUMAN	P36941 homo sapien
2	51	47.2	897	1 AAC2_CHICK	P20111 gallus gall
3	49	45.4	172	1 VPI9_CRV	P17457 cymbidium r
4	49	45.4	2142	1 BAT2_HUMAN	P48634 homo sapien
5	48	44.4	568	1 RGS3_MOUSE	Q9dc04 mus musculus
6	47	43.5	147	1 YAU8_SCHPO	Q10163 schizosacch
7	47	43.5	340	1 HEM6_SYNY3	P72848 synchocyst
8	46	42.6	173	1 VPI9_CNV	P15184 cucumber ne
9	46	42.6	1124	1 JAK3_HUMAN	P52333 homo sapien
10	45	41.7	415	1 TNR3_MOUSE	P50284 mus musculus
11	44	40.7	477	1 DXR_ARATH	Q9xf59 arabidopsis
12	44	40.7	502	1 RM02_ORYSA	P92812 oryza sativ
13	44	40.7	584	1 VATA_METHH	O27036 methanobact
14	44	40.7	1132	1 BAT3_HUMAN	P46379 homo sapien
15	43.5	40.3	1733	1 VNDA_PPVKA	P33485 pseudorabie
16	43.5	40.3	2364	1 PGCA_BOVIN	P13608 bos taurus
17	43	39.8	165	1 RS10_MOUSE	P09900 mus musculus
18	43	39.8	451	1 MB31_ARATH	O04309 arabidopsis
19	43	39.8	788	1 CY14_NEUCR	P23622 neurospora
20	43	39.8	894	1 AAC2_HUMAN	P35609 homo sapien
21	43	39.8	894	1 AAC2_MOUSE	Q9ji91 mus musculus
22	43	39.8	1213	1 FMN_CHICK	Q05858 gallus gall
23	42	38.9	166	1 REL1_PANTR	P51454 pan troglod
24	42	38.9	185	1 REL1_HUMAN	P04808 homo sapien
25	42	38.9	193	1 TNR3_HUMAN	P32370 homo sapien
26	42	38.9	208	1 YST2_SYNP7	P33772 synchococc
27	42	38.9	418	1 CDL5_HUMAN	Q14004 homo sapien
28	42	38.9	436	1 RHO_AQUAE	O67031 aquifex aeo
29	42	38.9	656	1 DNAA_STRCO	P27902 streptomyce
30	42	38.9	733	1 NIBL_HUMAN	Q96tal homo sapien
31	42	38.9	752	1 CO2_HUMAN	P06681 homo sapien
32	42	38.9	826	1 RGS5_HUMAN	O43374 homo sapien
33	42	38.9	894	1 MTP_HUMAN	P55157 homo sapien

34	42	38.9	1075	1 CNRA_ALCEU	P37372 alcaligenes
35	42	38.9	1076	1 NCCA_ALCXX	Q44586 alcaligenes
36	42	38.9	1100	1 JAK3_RAT	Q63272 rattus norv
37	42	38.9	1299	1 JAK3_MOUSE	O62137 mus musculu
38	42	38.9	2774	1 MAPA_RAT	P34326 rattus norv
39	41.5	38.4	487	1 Y442_MYCTU	Q9wvs8 mus musculu
40	41.5	38.4	806	1 MK07_MOUSE	Q13164 homo sapien
41	41.5	38.4	815	1 MK07_HUMAN	Q04637 homo sapien
42	41.5	38.4	1395	1 IF4G_HUMAN	O28343 canis famil
43	41.5	38.4	2333	1 PGCA_CANFA	O77302 lumbricus r
44	41	38.0	156	1 RS10_LUMRU	P14353 human spuma
45	41	38.0	205	1 BEL1_FOAMY	

ALIGNMENTS

RESULT 1	TNR3_HUMAN	STANDARD;	PRT;	435 AA.
ID	TNR3_HUMAN	STANDARD;	PRT;	435 AA.
AC	P36941:			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Tumor necrosis factor receptor superfamily member 3 precursor			
DE	(lymphotoxin-beta receptor) (Tumor necrosis factor receptor 2 related			
DE	protein) (Tumor necrosis factor C receptor).			
GN	LTBR OR TNFRSF3 OR TNFCR.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=93252381; PubMed=8486360;			
RA	Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;			
RT	"Construction and evaluation of a hncDNA library of human 12p			
RT	transcribed sequences derived from a somatic cell hybrid.;			
RL	Genomics 16:214-218(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung;			
RA	Strausberg R.;			
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
RP	FUNCTION.			
RX	MEDLINE=94225209; PubMed=8171323;			
RA	Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,			
RA	Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;			
RT	"A lymphotoxin-beta-specific receptor.;"			
RL	Science 264:707-710(1994).			
RN	[4]			
RP	CHARACTERIZATION.			
RX	MEDLINE=99223511; PubMed=10207006;			
RA	Wu M.-Y., Wang P.-Y., Han S.-H., Hsieh S.-L.;			
RT	"The cytoplasmic domain of the lymphotoxin-beta receptor mediates cell			
RT	death in HeLa cells.;"			
RL	J. Biol. Chem. 274:11868-11873(1999).			
RN	[5]			
RP	FUNCTION.			
RX	MEDLINE=20261554; PubMed=10799510;			
RA	Rooney I.A., Butrovich K.D., Glass A.A., Borboroglu S., Benedict C.A.,			
RA	Whitbeck J.C., Cohen G.H., Eisenberg R.J., Ware C.F.;			
RT	"The lymphotoxin-beta receptor is necessary and sufficient for			
RT	LIGHT-mediated apoptosis of tumor cells.;"			
RL	J. Biol. Chem. 275:14307-14315(2000).			
CC	-!- FUNCTION: Receptor for the heterotrimeric lymphotoxin containing			
CC	LTA and LTB, and for TNFSF14/LIGHT. Promotes apoptosis via TRAF3			
CC	and TRAF5. May play a role in the development of lymphoid organs.			
CC	-!- SUBUNIT: Self-associates.			
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-!- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.			

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L04270; AAA36757.1; -;
 DR EMBL; BC026262; AAH26262.1; -;
 DR HSSP; P25942; ICDP.
 DR Genew; HGNC:6718; LTBR.
 DR MIM; 600979; -;
 DR InterPro; IPR001368; TNFR_c6.
 DR Pfam; PF00020; TNFR_c6; 4.
 DR ProDom; PD000771; TNFR_c6; 1.
 DR SMART; SM00208; TNFR; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS00050; TNFR_NGFR_2; 3.
 KW Receptor; Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 435 TUMOR NECROSIS FACTOR RECEPTOR
 FT DOMAIN 31 227 SUPRACELLULAR MEMBER 3.
 FT TRANSMEM 228 248 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 249 435 POTENTIAL.
 FT REPEAT 42 81 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 82 124 TNFR-CYS 1.
 FT REPEAT 125 168 TNFR-CYS 2.
 FT REPEAT 169 211 TNFR-CYS 3.
 FT REPEAT 212 254 TNFR-CYS 4.
 FT DISULFID 43 58 BY SIMILARITY.
 FT DISULFID 59 72 BY SIMILARITY.
 FT DISULFID 73 86 BY SIMILARITY.
 FT DISULFID 87 100 BY SIMILARITY.
 FT DISULFID 101 116 BY SIMILARITY.
 FT DISULFID 117 132 BY SIMILARITY.
 FT DISULFID 133 148 BY SIMILARITY.
 FT DISULFID 149 165 BY SIMILARITY.
 FT DISULFID 166 185 BY SIMILARITY.
 FT CARBOHYD 40 40 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 177 177 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 435 AA; 46709 MW; 624626E6022F656F CRC64;

Query Match 74.1%; Score 80; DB 1; Length 435;
 Best Local Similarity 35.7%; Pred. No. 0.00019;
 Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 PLPPMS-----GSLKRRPQGE 20
 DB 216 PLPPMSGTLMVLAVLLPLAFLLLATVFCIWKSHPSLCRLKLSLLKRRPQGE 271
 |||||
 |||||

RESULT 2
 AAC2_CHICK STANDARD; PRT; 897 AA.
 AC P20111;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Alpha-actinin 2 (Alpha actinin skeletal muscle isoform 2) (F-actin
 DE cross linking protein).
 GN ACTN2.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89064821; PubMed=3197725;
 RA Arimura C., Suzuki T., Yanagisawa M., Inamura M., Hamada Y.,

Masaki T.;
 "Primary structure of chicken skeletal muscle and fibroblast alpha-
 actinins deduced from cDNA sequences.";
 Eur. J. Biochem. 177:649-655(1988).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=92070385; PubMed=1720388;
 RX Tokue Y., Goto S., Inamura M., Obinata T., Masaki T., Endo T.;
 "Transfection of chicken skeletal muscle alpha-actinin cDNA into
 nonmuscle and myogenic cells: dimerization is not essential for
 alpha-actinin to bind to microfilaments.";
 Exp. Cell Res. 197:158-167(1991).
 CC Exp. Cell Res. 197:158-167(1991).
 CC FUNCTION: F-ACTIN CROSS-LINKING PROTEIN WHICH IS THOUGHT TO ANCHOR
 ACTIN TO A VARIETY OF INTRACELLULAR STRUCTURES. THIS IS A BUNDLING
 PROTEIN.
 CC -1- SUBUNIT: HOMODIMER, ANTIPARALLEL.
 CC -1- SIMILARITY: CONTAINS 1 ACTIN-BINDING DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 ACTIN-BINDING (CH) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC -1- SIMILARITY: CONTAINS 4 SPECTRIN REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X13874; CAA32078.1; -;
 DR EMBL; X59247; CAA41935.1; -;
 DR PIR; S02032; S02032.
 DR PIR; S15481; S15481.
 DR HSSP; Q01082; 1BKR.
 DR InterPro; IPR001589; Actbind_actnin.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF00036; efhand; 2.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00435; spectrin; 4.
 DR ProDom; PD000012; EF-hand; 1.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00054; EFh; 2.
 DR SMART; SM00150; SPEC; 2.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS00018; EF_HAND; FALSE_NEG.
 KW Actin-binding; Calcium-binding; Repeat; Multigene family.
 FT DOMAIN 1 257 ACTIN-BINDING.
 FT DOMAIN 41 145 CH 1.
 FT DOMAIN 154 257 CH 2.
 FT REPEAT 284 394 SPECTRIN 1.
 FT REPEAT 404 509 SPECTRIN 2.
 FT REPEAT 519 630 SPECTRIN 3.
 FT REPEAT 640 743 SPECTRIN 4.
 FT CA_BIND 769 780 EF-HAND 1 (POTENTIAL).
 FT CA_BIND 805 816 EF-HAND 2 (POTENTIAL).
 SQ SEQUENCE 897 AA; 104275 MW; F4FAC12F7F4C8634 CRC64;
 Query Match 47.2%; Score 51; DB 1; Length 897;
 Best Local Similarity 52.6%; Pred. No. 7.6;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 2 LPPEMSGSLKRRPQGE 20
 DB 857 LPPEQAQYCIKRMPOYTG 875
 ||||| : |||||
 ||||| : |||||
 RESULT 3
 VP19_CRV STANDARD; PRT; 172 AA.
 ID VP19_CRV

AC P17457;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE Core protein P19.
 OS Cymbidium ringspot virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
 OC Tombusvirus.
 OX NCBI_TaxID=12144;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=89381709; PubMed=2778443;
 RX Grieco F., Burgyan J., Russo M.;
 RA "Nucleotide sequence of the 3'-terminal region of cymbidium ringspot
 RT virus RNA";
 RL J. Gen. Virol. 70:2533-2538(1989).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=89366663; PubMed=2771646;
 RX Grieco F., Burgyan J., Russo M.;
 RA "The nucleotide sequence of cymbidium ringspot virus RNA";
 RT Nucleic Acids Res. 17:6383-6383(1989).
 RL
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X15511; CAA33535.1; -
 DR PIR; JS0271; NKVGD2.
 DR PIR; S05459; S05459.
 DR InterPro; IPR004905; Tombus_P19.
 DR Pfam; PF03220; Tombus_P19; 1.
 KW Core protein.
 SQ SEQUENCE 172 AA; 19385 MW; F1069F14366D5D48 CRC64;
 Query Match 45.4%; Score 49; DB 1; Length 172;
 Best Local Similarity 47.1%; Pred. No. 2.6;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 PLPPMSGSLKRRPQG 17
 | | | | | : : : : :
 Db 147 PTPSERGNVRRRPEG 163
 RESULT 4
 ID BAT2_HUMAN STANDARD; PRT; 2142 AA.
 AC P48634;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Large proline-rich protein BAT2 (HLA-B-associated transcript 2).
 GN BAT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=T-cell.
 RX MEDLINE=90192810; PubMed=2156268;
 RA Banerji J., Sands J., Strominger J.L., Spies T.;
 RT "A gene pair from the human major histocompatibility complex encodes
 RT large proline-rich proteins with multiple repeated motifs and a
 RT single ubiquitin-like domain";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378(1990).
 RN [2]
 RN SEQUENCE OF 1-1860 FROM N.A.
 RX MEDLINE=93272029; PubMed=8499947;

RA Iris F.J.M., Bouqueleret L., Prieur S., Caterina D., Primas G.,
 RA Perrot V., Jurka J., Rodriguez-Tome P., Claverie J.-M., Dausset J.,
 RA Cohen D.;
 RT "Dense Alu clustering and a potential new member of the NF kappa B
 RT family within a 90 kilobase HLA class III segment";
 RL Nat. Genet. 3:137-145(1993).
 CC 1- FUNCTION: UNKNOWN.
 CC 1- TISSUE SPECIFICITY: LIMITED TO CELL-LINES OF LEUKEMIC ORIGIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M33509; AAA35585.1; -
 DR EMBL; M33518; AAA35586.1; -
 DR EMBL; M33512; AAA35586.1; JOINED.
 DR EMBL; Z15025; CAA78744.1; -
 DR PIR; B35098; B35098.
 DR PIR; S36152; S36152.
 DR Genew; HGNC:13918; BAT2.
 DR MIM; 142580; -
 KW Repeat.
 DR DOMAIN 519 524 POLY-PRO.
 FT DOMAIN 636 657 GLN-RICH.
 FT DOMAIN 684 688 POLY-PRO.
 FT DOMAIN 699 704 POLY-PRO.
 FT DOMAIN 814 821 POLY-PRO.
 FT DOMAIN 1340 1345 POLY-GLY.
 FT DOMAIN 1398 1403 POLY-GLY.
 FT DOMAIN 1436 1442 POLY-PRO.
 FT DOMAIN 1982 1991 POLY-PRO.
 FT DOMAIN 41 1795 4 X 57 AA TYPE A REPEATS.
 FT REPEAT 41 95 1-1.
 FT REPEAT 98 154 1-2.
 FT REPEAT 281 337 1-3.
 FT REPEAT 1740 1795 1-4.
 FT DOMAIN 337 549 2 X TYPE B REPEATS.
 FT REPEAT 337 418 2-1.
 FT REPEAT 476 549 2-2.
 FT DOMAIN 1899 2089 3 X 50 AA TYPE C REPEATS.
 FT REPEAT 1899 1948 3-1.
 FT REPEAT 1965 2014 3-2.
 FT REPEAT 2040 2089 3-3.
 FT CONFLICT 57 57 R -> A (IN REF. 2).
 FT CONFLICT 109 109 Q -> S (IN REF. 2).
 FT CONFLICT 414 414 P -> PPHRGAGNMGPP (IN REF. 2).
 FT CONFLICT 532 532 T -> K (IN REF. 2).
 FT CONFLICT 682 682 Q -> K (IN REF. 2).
 FT CONFLICT 730 730 E -> D (IN REF. 2).
 FT CONFLICT 750 750 L -> R (IN REF. 2).
 FT CONFLICT 834 834 A -> T (IN REF. 2).
 FT CONFLICT 1035 1035 G -> A (IN REF. 2).
 FT CONFLICT 1068 1068 M -> L (IN REF. 2).
 FT CONFLICT 1285 1285 P -> R (IN REF. 2).
 FT CONFLICT 1400 1400 G -> A (IN REF. 2).
 FT CONFLICT 1611 1611 T -> S (IN REF. 2).
 FT CONFLICT 1729 1729 G -> A (IN REF. 2).
 SQ SEQUENCE 2142 AA; 227840 MW; 32DDF16B9B52420A CRC64;
 Query Match 45.4%; Score 49; DB 1; Length 2142;
 Best Local Similarity 47.4%; Pred. No. 38;
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 2 LPPEMSGSLKRRPQGE 20
 | | | | | : : : : :
 Db 1561 LPPHSSGFLGSRPEGP 1579

RESULT 5

FT	DOMAIN	443	559	RGS.
FT	VARSPLIC	1	MISSING (IN SHORT ISOFORM).	
FT	VARSPLIC	376	KMSGDLTCTEDSRRKSKNI -> MLRGMVLTNRGNLOR	
FT	SEQUENCE	568 AA;	61558 MW;	91939E2F3DA9F5C CRC64;
ST	SEQUENCE	568 AA;	61558 MW;	91939E2F3DA9F5C CRC64;
Query Match		44.4%;	Score 48;	DB 1; Length 568;
Best Local Similarity		50.0%;	Pred. No. 13;	
Matches		10;	Conservative	2; Mismatches
QY	1 PLPPPMGSLKKRRPQGEGP	20		
ID	YAUS_SCHPO	STANDARD:	PRT:	147 RA.
AC	Q10163;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DT	Hypochemical protein C26A3.08 in chromosome I.			
DE	SPAC36A3.08.			
GN	Schizosaccharomyces pombe (Fission yeast).			
OS	Schizosaccharomycetes			
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;			
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;			
OC	Schizosaccharomycetes.			
OX	NCBI_TaxID=4896;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=972;			
RC	MEDLINE=21848401; PubMed=11859360;			
RA	Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,			
RA	Scoutros J., Peat N., Hayles J.J., Baker S., Basham D., Bowman S.,			
RA	Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,			
RA	Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,			
RA	Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson K.,			
RA	Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,			
RA	James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,			
RA	Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,			
RA	Olliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,			
RA	Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,			
RA	Skelton J., Simmonds M., Squares R., Walters S.V., Warren T., Whitehead S.,			
RA	Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,			
RA	Woodward J., Volkhardt G., Aert R., Robben J., Grymonprez B.,			
RA	Weitjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,			
RA	Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hubert H.,			
RA	Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,			
RA	Eger P., Zimmermann W., Wedler H., Wambutt R., Mottier S.,			
RA	Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,			
RA	Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,			
RA	Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,			
RA	Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,			
RA	Domiguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,			
RA	Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,			
RA	Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;			
RT	"The genome sequence of Schizosaccharomycetes pombe."			
RL	Nature 415:871-880(2002).			
CC	- SIMILARITY: BELONGS TO THE SNRNP SMB/SMN FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; Z69240; CA93231.1;			
DR	HSSP; PI4678; LD3B.			
DR	InterPro; IPRO01163; snRNP_Sm.			
DR	Pfam; PF01423; Sm; 1.			
KW	Signal transduction inhibitor; Alternative splicing; Phosphorylation.			


```

KW Hypothetical protein; Nuclear protein; Ribonucleoprotein; RNA-binding.
SQ SEQUENCE 147 AA; 15476 MW; 09FD4F2D583F0FF5 CRC64;
  Query Match 43.5%; Score 47; DB 1; Length 147;
  Best Local Similarity 47.8%; Pred. No. 4.3;
  Matches 11; Conservative 1; Mismatches 7; Indels 4; Gaps 1;

QY 1 PLPPMSGSLK----RRPQEG 19
   I I I I I I I I I I
Db 85 PMDPSMRGSLSGGVARPAG 107

RESULT 7
HEM6_SYNY3
ID HEM6_SYNY3 STANDARD; PRT; 340 AA.
AC P72848;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Coproporphyrinogen III oxidase, aerobic (EC 1.3.3.3)
DE (Coproporphyrinogenase) (Coprogen oxidase).
DE HEMF OR SL1185
GN Hemf OR SL1185
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
PI 1
SEQUENCE FROM N.A.
EMBL; M25270; AAA42906.1; -
DR InterPro; IPR001260; Coprogen_oxidase.
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
RA Okumura S., Shimizu S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RA "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -!- CATALYTIC ACTIVITY: Coproporphyrinogen-III + O(2) =
CC protoporphyrinogen-IX + 2 CO(2).
CC -!- COFACTOR: IRON (BY SIMILARITY).
CC -!- PATHWAY: Porphyrin biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE AEROBIC COPROPORPHYRINOGEN III OXIDASE
CC FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; D90901; BAA16863.1; -
DR InterPro; IPR001260; Coprogen_oxidase.
DR Pfam; PF01218; Coprogen_oxidase; 1.
DR PRINTS; PR00073; COPROGNOXDASE.
DR PROSITE; PS01021; COPROGEN OXIDASE; 1.
KW Porphyrin biosynthesis; Oxidoreductase; Iron; Complete proteome.
SQ SEQUENCE 340 AA; 38937 MW; 915623BD89587FD CRC64;
  Query Match 43.5%; Score 47; DB 1; Length 340;
  Best Local Similarity 55.6%; Pred. No. 11;
  Matches 10; Conservative 3; Mismatches 1; Indels 4; Gaps 1;

QY 2 LPPEMSGSLKRRPQEG 19
   I I I I I I I I I I
Db 87 LPP-----SILKORPEAEG 100

RESULT 8
VP19_CNV
ID VP19_CNV STANDARD; PRT; 173 AA.

```

```

AC PI5184;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE Core protein P19 (P20).
OS Cucurbit necrosis virus (CNV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tombusviridae;
OC Tombusvirus.
OX NCBI_TaxID=12143;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89204896; PubMed=2705296;
RA Rochon D.M., Tremaine J.H.;
RT "Complete nucleotide sequence of the cucumber necrosis virus genome.";
RL Virology 169:251-259(1989).
CC -!- SIMILARITY: TO OTHER TOMBUSVIRUSES CORE PROTEIN P19.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M25270; AAA42906.1; -
DR PIR; JA0133; NKVGCU.
DR InterPro; IPR004905; Tombus_P19.
DR Pfam; PF03220; Tombus_P19; 1.
KW Core protein.
SQ SEQUENCE 173 AA; 19808 MW; D05F870D4AA4C80F CRC64;
  Query Match 42.6%; Score 46; DB 1; Length 173;
  Best Local Similarity 53.3%; Pred. No. 7.2;
  Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 PPEMSGSLKRRPQ 17
   I I I I I I I I I I
Db 149 PCEVGNVSRRRPQ 163

RESULT 9
JAK3_HUMAN
ID JAK3_HUMAN STANDARD; PRT; 1124 AA.
AC P52333; Q13259; Q13260; Q13611;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase JAK3 (EC 2.7.1.112) (Janus kinase 3) (JAK-3)
DE (Leukocyte janus kinase) (L-JAK).
GN JAK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=94294384; PubMed=802790;
RA Kawamura M., McVicar D.W., Johnston J.A., Blake T.B., Chen Y.-Q.,
RA Lai B.K., Lloyd A.R., Kelvin D.J., Staples J.E., Ortaldo J.R.,
RA O'Shea J.J.;
RT "Molecular cloning of L-JAK, a Janus family protein-tyrosine kinase
RT expressed in natural killer cells and activated leukocytes.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:6374-6378(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RX MEDLINE=96027605; PubMed=7559633;
RA Lai K.S., Jin Y., Graham D.K., Witthuhn B.A., Ihle J.N., Liu E.T.;
RT "A kinase-deficient splice variant of the human JAK3 is expressed in
RT hematopoietic and epithelial cancer cells.";
RL J. Biol. Chem. 270:25028-25036(1995).
RN [3]
RP SEQUENCE OF 36-191 FROM N.A.

```

CC	MEDLINE=96278845; PubMed=8662778;
CC	Verbsky J.W., Bach E.A., Fang Y.F., Yang L., Randolph D.A.,
CC	Fields L.E.;
CC	"Expression of Janus kinase 3 in human endothelial and other non-
CC	lymphoid and non-myeloid cells.";
CC	J. Biol. Chem. 271:13976-13980(1996).
CC	[4]
CC	VARIANT SCID CYS-100.
CC	MEDLINE=95388142; PubMed=7659163;
CC	Macchi P., Villa A., Gillani S., Sacco M.G., Frattini A., Porta F.,
CC	Ugazio A.G., Johnston J.A., Candotti F., O'Shea J.J., Vezzoni P.,
CC	Notarangelo L.D.;
CC	"Mutations of Jak-3 gene in patients with autosomal severe combined
CC	immune deficiency (SCID).";
CC	Nature 377:65-68(1995).
CC	[5]
CC	VARIANTS SCID GLY-481; 586-LEU-MET-592 DEL AND ARG-759.
CC	MEDLINE=98022793; PubMed=9354668;
CC	Candotti F., Oakes S.A., Johnston J.A., Gillani S., Schumacher R.F.,
CC	Mella P., Fiorini M., Ugazio A.G., Badolato R., Notarangelo L.D.,
CC	Bozzi F., Macchi P., Strina D., Vezzoni P., Blaese R.M., O'Shea J.J.,
CC	Villa A.;
CC	"Structural and functional basis for JAK3-deficient severe combined
CC	immunodeficiency.";
CC	Blood 90:3996-4003(1997).
CC	[6]
CC	VARIANT SCID TRP-582.
CC	MEDLINE=98423994; PubMed=9753072;
CC	Bozzi F., LeFranc G., Villa A., Badolato R., Schumacher R.F.,
CC	Khalil G., Loiselet J., Bresciani S., O'Shea J.J., Vezzoni P.,
CC	Notarangelo L.D., Candotti F.;
CC	"Molecular and biochemical characterization of JAK3 deficiency in a
CC	patient with severe combined immunodeficiency over 20 years after
CC	bone marrow transplantation: implications for treatment.";
CC	Br. J. Haematol. 102:1363-1366(1998).
CC	[7]
CC	VARIANTS SCID ARG-151; ILE-722 AND SER-910.
CC	MEDLINE=20435064; PubMed=10982185;
CC	Schumacher R.F., Mella P., Badolato R., Fiorini M., Savoldi G.,
CC	Gillani S., Villa A., Candotti F., Tampalini A., O'Shea J.J.,
CC	Notarangelo L.D.;
CC	"Complete genomic organization of the human JAK3 gene and mutation
CC	analysis in severe combined immunodeficiency by single-strand
CC	conformation polymorphism.";
CC	Hum. Genet. 106:773-79(2000).
CC	[-] FUNCTION: TYROSINE KINASE OF THE NON-RECEPTOR TYPE, INVOLVED IN
CC	THE INTERLEUKIN-2 AND INTERLEUKIN-4 SIGNALING PATHWAY.
CC	PHOSPHORYLATES STAT6, IRS1, IRG2 AND PI3K.
CC	[-] CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC	tyrosine phosphate.
CC	[-] SUBCELLULAR LOCATION: WHOLLY INTRACELLULAR, POSSIBLY MEMBRANE
CC	ASSOCIATED (BY SIMILARITY).
CC	[-] ALTERNATIVE PRODUCTS: 3 isoforms; 1/JAK3B/breast-JAK3,
CC	2/JAK3S/spleen-JAK3 (shown here) and 3/JAK3M/activated monocytes-
CC	JAK3; are produced by alternative splicing. Isoform 1 may be
CC	defective as it lack some part of the kinase domain.
CC	[-] TISSUE SPECIFICITY: IN NK CELLS AND AN NK-LIKE CELL LINE BUT NOT
CC	IN RESTING T CELLS OR IN OTHER TISSUES. THE S-FORM IS MORE
CC	COMMONLY SEEN IN HEMATOPOIETIC LINES, WHEREAS THE B- AND M-FORMS
CC	ARE DETECTED IN CELLS BOTH OF HEMATOPOIETIC AND EPITHELIAL
CC	ORGANS.
CC	[-] DOMAIN: POSSESSES TWO PHOSPHOTRANSFERASE DOMAINS. THE SECOND ONE
CC	PROBABLY CONTAINS THE CATALYTIC DOMAIN (BY SIMILARITY), WHILE THE
CC	PRESENCE OF SLIGHT DIFFERENCES SUGGEST A DIFFERENT ROLE FOR
CC	DOMAIN 1.
CC	[-] PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-2 AND IL-4.
CC	[-] DISEASE: DEFECTS IN JAK3 ARE A CAUSE OF RECESSIVE T-CELL
CC	NEGATIVE/B-C-CELL POSITIVE SEVERE COMBINED IMMUNODEFICIENCY (T-B+
CC	SCID). A CONDITION CHARACTERIZED BY THE ABSENCE OF CIRCULATING
CC	MATURE T-LYMPHOCYTES AND NK CELLS, NORMAL TO ELEVATED NUMBERS OF
CC	NONFUNCTIONAL B-LYMPHOCYTES, AND MARKED HYPOPLASIA OF LYMPHOID
CC	TISSUES.
CC	[-] REFERENCES TO THIS ENTRY FROM PROTEIN KINASES
CC	RA
CC	RX
CC	RA
CC	RT
CC	RL
CC	RN
CC	RP
CC	RR
CC	RY
CC	SA
CC	SB
CC	SC
CC	SD
CC	SE
CC	SF
CC	SG
CC	SH
CC	SI
CC	SJ
CC	SK
CC	SL
CC	SM
CC	SN
CC	SO
CC	SP
CC	SR
CC	SS
CC	ST
CC	SV
CC	SW
CC	SX
CC	SY
CC	SZ
CC	TA
CC	TB
CC	TC
CC	TD
CC	TE
CC	TF
CC	TH
CC	TI
CC	TJ
CC	TK
CC	TL
CC	TM
CC	TN
CC	TO
CC	TP
CC	TR
CC	TS
CC	TT
CC	TU
CC	TV
CC	TW
CC	TX
CC	TY
CC	TZ
CC	UA
CC	UB
CC	UC
CC	UD
CC	UE
CC	UF
CC	UG
CC	UH
CC	UI
CC	UJ
CC	UK
CC	UL
CC	UM
CC	UN
CC	UO
CC	UP
CC	UR
CC	US
CC	UT
CC	UU
CC	UV
CC	UX
CC	UY
CC	UZ
CC	VA
CC	VB
CC	VC
CC	VD
CC	VE
CC	VF
CC	VG
CC	VH
CC	VI
CC	VJ
CC	VK
CC	VL
CC	VM
CC	VN
CC	VO
CC	VP
CC	VQ
CC	VR
CC	VS
CC	VT
CC	VU
CC	VV</

FT	SIGNAL	1	30	POTENTIAL.
FT	CHAIN	31	415	TUMOR NECROSIS FACTOR RECEPTOR
FT				SUPERFAMILY MEMBER 3.
FT				EXTRACELLULAR (POTENTIAL).
FT				POTENTIAL.
FT	DOMAIN	31	223	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	224	244	
FT	DOMAIN	245	415	
FT	REPEAT	42	81	TNFR-CYS 1.
FT	REPEAT	82	124	TNFR-CYS 2.
FT	REPEAT	125	170	TNFR-CYS 3.
FT	REPEAT	171	213	TNFR-CYS 4.
FT	DISULFID	43	58	BY SIMILARITY.
FT	DISULFID	59	72	BY SIMILARITY.
FT	DISULFID	62	80	BY SIMILARITY.
FT	DISULFID	83	98	BY SIMILARITY.
FT	DISULFID	101	116	BY SIMILARITY.
FT	DISULFID	104	124	BY SIMILARITY.
FT	DISULFID	126	132	BY SIMILARITY.
FT	DISULFID	139	150	BY SIMILARITY.
FT	DISULFID	142	169	BY SIMILARITY.
FT	DISULFID	172	187	BY SIMILARITY.
FT	CARBOHYD	40	40	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	179	179	N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE	415 AA;	44956 MW;	29B326A56AEF661 CRC64;

Query Match	41.7%;	Score 45;	DB 1;	Length 415;
Best Local Similarity	72.7%;	Pred. No. 26;		
Matches 8;	Conservative	2;	Mismatches 1;	Indels 0; Gaps

QY	8	GSLLKRRPGE	18
	:	:	
DB	255	GTLLKRRPGE	265

RESULT 11			
DXR-ARATH	STANDARD;	PRT;	477 AA.
ID	DXR-ARATH		
AC	Q9XFS9; Q9M6U2;		
DT	30-MAY-2000 (Rel. 39, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	1-deoxy-D-xylulose 5-phosphate reductoisomerase, chloroplast precursor		
DE	(EC 1.1.1.267) (DXP reductoisomerase) (1-deoxyxylulose-5-phosphate		
DE	reductoisomerase).		
GN	DXR OR AT5G62790 OR MQB2.11.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
NCBI_Taxid=3702;			
LN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Columbia;		
RA	Campos N., Lois L.M., Cunillera N., Carretero L., Ahumada I.,		
RA	Hoeffler J.-F., Pale-Grosdemange C., Rohmer M., Ferrer A., Boronat A.,		
RT	"Isolation and characterization of a cDNA from Arabidopsis thaliana		
RT	encoding 1-deoxy-D-xylulose 5-phosphate reductoisomerase, the first		
RT	committed enzyme of the non-mevalonate pathway for isoprenoid		
RT	biosynthesis.";		
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Columbia;		
RX	MEDLINE=98290546; PubMed=9628582;		
RA	Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,		
RA	Tabata S.;		
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. IV.		
RT	Sequence features of the regions of 1,456,315 bp covered by nineteen		
RT	physically assigned pl and TAC clones.";		
RL	DNA Res. 5:41-54(1998).		
RP	[3]		
RP	SEQUENCE OF 72-477 FROM N.A.		
RC	STRAIN=cv. Columbia;		
RA	Schwender J., Mueller C., Zeidler J., Lichtenthaler H.K.;		

```

RT CC the European Bioinformatics Institute. There are no restrictions on its
RT CC use by non-profit institutions as long as its content is in no way
RL CC modified and this statement is not removed. Usage by and for commercial
RL CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; D78336; BAAL1350.1; ALT_SEQ.
CC CC HSP; P04257; 1RL2.
CC CC InterPro; IPR002171; Ribosomal_L2.
CC CC Pfam; PF00181; Ribosomal_L2; 1.
CC CC TIGRFAMs; TIGR01171; rplB_bact; 1.
CC CC PROSITE; PS00467; RIBOSOMAL_L2; 1.
CC CC Ribosomal protein; Mitochondrion; RNA editing.
CC CC KW SEQUENCE 502 AA; 55095 MW; 4ACAD1CD34A31F7D CRC64;
CC CC -----
CC CC Query Match 40.7%; Score 44; DB 1; Length 502;
CC CC Best Local Similarity 57.1%; Pred. No. 44;
CC CC Matches 12; Conservative 1; Mismatches 6; Indels 2; Gaps 1;
CC CC -----
CC CC QY 2 LPPEMS--GSLKRRPQGE 20
CC CC ||| ||| : ||||
CC CC Db 248 LPVEASRFGLPAKPPIGE 268
CC CC -----
CC CC RESULT 13
CC CC VATA_METTH STANDARD; PRT; 584 AA.
CC CC ID VATA_METTH
CC CC AC 027036;
CC CC DT 16-OCT-2001 (Rel. 40, Created)
CC CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC CC DE V-type ATP synthase alpha chain (EC 3.6.3.14) (V-type ATPase subunit
CC CC DE A).
CC CC DN ATPA OR MTH955.
CC CC OS Methanobacterium thermoautotrophicum.
CC CC OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
CC CC OC Methanobacteriaceae; Methanothermobacter.
CC CC OX NCBI_TaxID=187420;
CC CC RN [1]
CC CC RC SEQUENCE FROM N.A.
CC CC RX STRAIN=Delta H;
CC CC RA MEDLINE=98037514; PubMed=9371463;
CC CC RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
CC CC RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
CC CC RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
CC CC RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
CC CC RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
CC CC RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
CC CC RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
CC CC RT "Complete genome sequence of Methanobacterium thermoautotrophicum
CC CC RT deltaH: functional analysis and comparative genomics.";
CC CC RL J. Bacteriol. 179:7135-7155(1997).
CC CC CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC CC GRADIENT ACROSS THE MEMBRANE. THE ARCHAEAL ALPHA CHAIN IS A
CC CC CATALYTIC SUBUNIT.
CC CC CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) -> ADP + phosphate +
CC CC H(+)(Out).
CC CC CC -1- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AE000869; AAB85451.1;
CC CC DR InterPro; IPR000793; ATPase_a/bc.
CC CC DR InterPro; IPR004100; ATPase_a/bn.
CC CC DR InterPro; IPR000194; ATPase_a/bn.
CC CC DR Pfam; PF00006; ATP-synt_ab; 1.
CC CC DR Pfam; PF00306; ATP-synt_ab_C; 1.
CC CC -----

```

```

RT CC "Cloning and heterologous expression of a cDNA coding for 1-deoxy-d-
RT CC xylulose-5-phosphate reductoisomerase of Arabidopsis thaliana.";
RL CC Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
CC CC of 1-deoxy-D-xylulose-5-phosphate (DXP) to 2-C-methyl-D-erythritol
CC CC 4-phosphate (MEP) (By similarity).
CC CC CC -1- CATALYTIC ACTIVITY: 2-C-methyl-D-erythritol 4-phosphate + NADP(+)
CC CC = 1-deoxy-D-xylulose 5-phosphate + NADPH.
CC CC CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
CC CC step.
CC CC CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC CC CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AF148852; AAF73140.1; -.
CC CC DR EMBL; AB009053; BAB10848.1; -.
CC CC DR EMBL; AJ242588; CAB43344.1; -.
CC CC DR InterPro; IPR003821; DXR_reductoisom; 1.
CC CC DR Pfam; PF02670; DXR_reductoisom; 1.
CC CC DR TIGRFAMs; TIGR00243; DXR; 1.
CC CC KW Isoprene biosynthesis; Oxidoreductase; NADP; Chloroplast;
CC CC TRANSIT PEPTIDE.
CC CC FT CHAIN ? 477 CHLOROPLAST (POTENTIAL).
CC CC FT CHAIN ? 477 1-DEOXY-D-XYLULOSE 5-PHOSPHATE
CC CC FT CHAIN ? 477 REDUCTOISOMERASE.
CC CC FT NP_BIND 86 93 NADPH (POTENTIAL).
CC CC SEQUENCE 477 AA; 51963 MW; C06A455AB73ACA7C CRC64;
CC CC Query Match 40.7%; Score 44; DB 1; Length 477;
CC CC Best Local Similarity 47.4%; Pred. No. 42;
CC CC Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
CC CC -----
CC CC QY 1 PLPPMSGSLKRRPQGE 19
CC CC ||| ||| ||| |||
CC CC Db 25 PIPKLSGFSLRNQG 43
CC CC -----
CC CC RESULT 12
CC CC RM02_ORYSA STANDARD; PRT; 502 AA.
CC CC AC P2812;
CC CC DT 15-DEC-1998 (Rel. 37, Created)
CC CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
CC CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC CC DE Mitochondrial 60S ribosomal protein L2.
CC CC GN RPL2.
CC CC OS Oryza sativa (Rice).
CC CC OG Mitochondrion.
CC CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC CC OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
CC CC OC Ehrhartoideae; Oryzeae; Oryza.
CC CC OX NCBI_TaxID=4530;
CC CC RN [1]
CC CC RP STRAIN=cv. Nipponbare; TISSUE=Shoot;
CC CC RX MEDLINE=96400039; PubMed=8806415;
CC CC RA Kubo N., Ozawa K., Hino T., Kadowaki K.;
CC CC RT "A ribosomal protein L2 gene is transcribed, spliced, and edited at
CC CC one site in rice mitochondria.";
CC CC RL Plant Mol. Biol. 31:853-862(1996).
CC CC CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC CC CC -1- MISCELLANEOUS: POSITION 404 IS MODIFIED BY RNA EDITING.
CC CC CC -1- SIMILARITY: BELONGS TO THE L2P FAMILY OF RIBOSOMAL PROTEINS.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

DR Pfam; PF02874; ATP-synt_ab_N; 1.
 DR TIGRfams; TIGR01043; ATP_syn_A_arch; 1.
 DR PROSITE; PS00152; ATPase.ALPHA.BETA; 1.
 KW Hydrolyase; ATP synthesis; Hydrogen ion transport; ATP-binding;
 KW Complete proteome. 240 ATP (POTENTIAL).
 FT NP_BIND 233
 SQ SEQUENCE 584 AA; 64894 MW; 90F01BD38B3D6D78 CRC64;

Query Match 40.7%; Score 44; DB 1; Length 584;
 Best Local Similarity 47.1%; Pred. No. 52;
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 LPPMSGSLKRRPQGE 18
 :||:|||||
 Db 151 IPPNVEGKLTWIA PQGE 167

RESULT 14
 BAT3_HUMAN STANDARD; PRT; 1132 AA.
 ID BAT3_HUMAN
 AC P46379;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Large proline-rich protein BAT3 (HLA-B-associated transcript 3).
 GN BAT3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=T-cell;
 RX MEDLINE=90192810; PubMed=2156268;
 RA Banerji J., Sands J., Strominger J.L., Spies T.;
 RT "A gene pair from the human major histocompatibility complex encodes
 RT large proline-rich proteins with multiple repeated motifs and a
 RT single ubiquitin-like domain."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378(1990).
 CC 1- FUNCTION: UNKNOWN
 CC 1- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M33519; AAA35587.1; -
 DR EMBL; M33521; AAA35588.1; -
 DR EMBL; M33520; AAA35588.1; JOINED.
 DR HSSP; P02248; IUBI.
 DR Genew; HGNC:13919; BAT3.
 DR MIM; 142590; -
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00213; UBQ; 1.
 DR PROSITE; PS00299; UBIQUITIN_1; 1.
 DR PROSITE; PS50053; UBIQUITIN_2; 1.
 DR Repeat.
 KW DOMAIN 17 77 UBIQUITIN-LIKE.
 FT DOMAIN 202 207 POLY-PRO.
 FT DOMAIN 242 636 4 X 29 AA APPROXIMATE REPEATS.
 FT REPEAT 242 270 1.
 FT REPEAT 415 443 2.
 FT REPEAT 574 602 3.
 FT REPEAT 608 636 4.
 FT DOMAIN 657 670 POLY-PRO.
 SQ SEQUENCE 1132 AA; 119504 MW; E28CA8A78C38DD18 CRC64;

Query Match 40.7%; Score 44; DB 1; Length 1132;

Best Local Similarity 50.0%; Pred. No. 1e+02;
 Matches 13; Conservative 1; Mismatches 6; Indels 6; Gaps 2;

QY 1 PLPP-----EMSGSLKRRP--QGE GP 20
 :||:|||||
 Db 1034 POPPLSDAYLSGMPAKRRKTKMQGEGP 1059

RESULT 15
 VNUA_PVKA STANDARD; PRT; 1733 AA.
 ID VNUA_PVKA
 AC P33485;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Probable nuclear antigen.
 OS Pseudorabies virus (strain Kaplan) (PRV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=33703;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91021039; PubMed=2171211;
 RA Vitek C., Kozmik Z., Paces V., Schirm S., Schwyzer M.;
 RT "Pseudorabies virus immediate-early gene overlaps with an oppositely
 RT oriented open reading frame: characterization of their promoter and
 RT enhancer regions".
 RL Virology 179:365-377(1990).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M34651; AAA47471.1; -
 DR PIR; B45344; B45344.
 FT DOMAIN 112 117 POLY-THR.
 FT DOMAIN 179 1733 GLY-RICH.
 FT DOMAIN 192 196 POLY-SER.
 FT DOMAIN 271 298 POLY-PRO.
 FT DOMAIN 304 308 POLY-ARG.
 FT DOMAIN 883 889 POLY-GLY.
 FT DOMAIN 1398 1405 POLY-GLY.
 SQ SEQUENCE 1733 AA; 172166 MW; 0C8CD8BE475BB5E2 CRC64;

Query Match 40.3%; Score 43.5; DB 1; Length 1733;
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;
 Matches 10; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1 PLPPMSGSLKRRPQGE GP 20
 :||:|||||
 Db 294 POPPPAGGS-ARRRRGGGP 312

Search completed: April 8, 2003, 10:49:43
 Job time : 4.38425 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:46:30 ; Search time 2.52983 Seconds
(without alignments) 760.006 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPMSGSLKRRPQGECP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	74.1	435	I54182	tumor necrosis fac
2	51	47.2	897	S02032	alpha-actinin 2, s
3	50	46.3	125	E72716	hypothetical prote
4	49	45.4	172	NKVGD2	core protein p19 -
5	49	45.4	1460	1 EDSEIF	immediate-early pr
6	49	45.4	1870	2 S37671	MHC class III hist
7	49	45.4	1872	2 S36152	MHC class III hist
8	49	45.4	2142	2 B35098	MHC class III hist
9	47	43.5	147	2 T38396	small nuclear ribo
10	47	43.5	340	2 S74712	coproporphyrinogen
11	46.5	43.1	138	2 B84430	hypothetical prote
12	46.5	43.1	572	2 T45392	hypothetical prote
13	46	42.6	173	1 NKVVCU	core protein p20 -
14	46	42.6	471	2 A84741	probable myosins
15	46	42.6	501	2 S76563	hypothetical prote
16	46	42.6	528	2 T52092	DNA-binding protei
17	46	42.6	1124	2 A55747	L-JAK protein-tyro
18	45.5	42.1	363	2 A56940	integral membrane
19	45	41.7	132	2 AE2344	hypothetical prote
20	45	41.7	210	2 T36862	probable two-compo
21	45	41.7	583	2 E75529	probable peptide A
22	45	41.7	702	2 H90757	probable oxidoredu
23	45	41.7	702	2 F85621	probable oxidoredu
24	45	41.7	2938	2 T30249	cell proliferation
25	44.5	41.2	596	2 F84589	probable protein k
26	44.5	41.2	3198	2 A43426	collagen alpha 2 f
27	44	40.7	472	2 G71503	probable replicat
28	44	40.7	472	2 H81665	replicative DNA he
29	44	40.7	502	2 T03019	probable ribosomal

30	44	40.7	505	2 C98300	hypothetical prote
31	44	40.7	505	2 AD2983	hypothetical prote
32	44	40.7	584	2 G69227	ATP synthase, subu
33	44	40.7	803	2 F59433	RhGAP protein [lm
34	44	40.7	1132	2 A35098	MHC class III hist
35	43.5	40.3	1733	1 B45344	probable nuclear a
36	43.5	40.3	2327	2 T42630	aggreccan - bovine
37	43	39.8	165	1 R3RT10	ribosomal protein
38	43	39.8	165	2 H59404	ribosomal protein
39	43	39.8	284	2 AD3275	beta-ureidopropion
40	43	39.8	497	2 B75218	glycerol-3-phospha
41	43	39.8	696	2 S74984	long-chain-fatty-a
42	43	39.8	781	2 A37956	sulfate permease I
43	43	39.8	894	1 FAHUA2	alpha-actinin 2 -
44	43	39.8	1213	2 A41724	limb deformity (ld
45	43	39.8	1531	2 T42218	slit-1 protein hom

ALIGNMENTS

RESULT 1

I54182
tumor necrosis factor receptor 2-related protein - human
C:Species: Homo sapiens (man)
C>Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 17-Mar-2000
C:Accession: I54182
R:Baens, M.; Chaffanet, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.
Genomics 16, 214-218, 1993
A:Title: Construction and evaluation of a hncDNA library of human 12p transcribed seq
A:Reference number: I54182; MUID:93252381; PMID:8486360
A:Accession: I54182
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-435 <RES>
A:Cross-references: GB:L04270; NID:G339761; PIDN:AAA36757.1; PID:G339762
C:Genetics:
A:Gene: GDB:LTBR
A:Cross-references: GDB:1230195; OMIM:600979
A:Map position: 12p13.3-12p13.1
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology

Query Match 74.1%; Score 80; DB 2; Length 435;
Best Local Similarity 35.7%; Pred. No. 0.00037;
Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY	1	PLPPMS-	-----GSLKRRPQGECP 20
DB	216	PLPPMSGTMLAVLLPLAFFLLATVFCIWKSHPSLCRLKLSLLKRRPQGECP 271	

RESULT 2

S02032
alpha-actinin 2, skeletal muscle splice form SK - chicken
C:Species: Gallus gallus (chicken)
C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 13-Aug-1999
C:Accession: S02032; I50604; S36393; S27391; S15481
R:Arimura, C.; Suzuki, T.; Yanagisawa, M.; Imamura, M.; Hamada, Y.; Masaki, T.
Eur. J. Biochem. 177, 649-655, 1988
A:Title: Primary structure of chicken skeletal muscle and fibroblast alpha-actinins d
A:Reference number: S02032; MUID:89064821; PMID:3197725
A:Accession: S02032
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-897 <ARI>
A:Cross-references: GB:X13874; NID:G63787; PIDN:CAA32078.1; PID:G63788
A:Note: part of this sequence was confirmed by protein sequencing
R:Tokue, Y.; Goto, S.; Imamura, M.; Obinata, T.; Masaki, T.; Endo, T.
Exp. Cell Res. 197, 158-167, 1991
A:Title: Transfection of chicken skeletal muscle alpha-actinin cDNA into nonmuscle an
A:Reference number: I50604; MUID:92070385; PMID:1720388
A:Accession: I50604
A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA
A:Residues: 1-897 <VO>
A:A:Cross-references: EMBL:X59247; MID:g63030; PIDN:CAA41935.1; PID:g63031
R:Par, T.
A:Reference number: S36393
A:Accession: S36393
A:Molecule type: DNA
A:Residues: 662-867 <PAR1>
A:Cross-references: EMBL:X68801
R:Par, T.; Waites, G.T.; Patel, B.; Millake, D.B.; Critchley, D.R.
Eur. J. Biochem. 210, 801-809, 1992
A:Title: A chick skeletal-muscle alpha-actinin gene gives rise to two al
A:Reference number: S27391; MUID:93130905; PMID:1483465
A:Accession: S27391
A:Molecule type: DNA
A:Residues: 707-867 <PAR2>
A:Cross-references: EMBL:X68801
A:Note: The authors translated the codon AAA for residue 714 as Leu and
C:Comment: The EF hand structures are predicted to be incapable of bindi
C:Comment: The amino-terminal domain is involved in actin binding, the m
nt chain.
C:Genetics:
A:Introns: 721/3; 770/3; 792/3; 845/3
A:Note: The list of introns is incomplete
C:Complex: homodimer
C:Superfamily: alpha-actinin; alpha-actinin actin-binding domain homolog
C:Keywords: actin binding; alternative splicing; duplication; EF hand; F
F:40-753/Domain: alpha-actinin actin-binding domain homology <ACT>
F:283-394/Domain: spectrin/dystrophin repeat homology <SP1>
F:403-509/Domain: spectrin/dystrophin repeat homology <SP2>
F:518-630/Domain: spectrin/dystrophin repeat homology <SP3>
F:639-743/Domain: spectrin/dystrophin repeat homology <SP4>
F:736-788/Domain: spectrin/dystrophin repeat homology <EF1>
F:792-824/Domain: calmodulin repeat homology <EF2>

Query Match 47.2%; Score 51; DB 2; Length 897;
Best Local Similarity 52.6%; Pred. No. 16;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps

QY 2 LPPEMSGSLKKRRPQGE 20
|||||: ||||| ||
DB 857 LPPEQAQYCKRNPPQVTGP 875

RESULT 3
E72716
hypothetical protein APE1148 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-
C:Accession: E72716
R:Kawabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.;
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.;
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Cre
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: E72716
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <KAW>
A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA80133.1; PID:
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1148

	Query Match	Best Local Similarity	Pruned	Score
QY	1	PLPPMSGV--SLKKRRPQGGP	20	46.38;
		: : :		54.58;
db	10	PLLPQDDGRPRLLRRGPGGGP	31	2;
		Matches	12;	Conservative

A: Molecule type: mRNA
A: Residues: 1-897 <T02>
A: Cross-references: EMBL:X59247; NID:g63030; PIDN:CAA41935.1; PID:g63031
R: Parr, T.
A: Reference number: 1992
A: Accession: S36393
A: Molecule type: DNA
A: Residues: 662-867 <P>
A: Cross-references: EMBL:X68801
R: Parr, T.; Waites, G.T.; Patel, B.; Millake, D.B.; Critchley, D.R.
Eur. J. Biochem. 210, 801-809, 1992
A: Title: A chick skeletal-muscle alpha-actinin gene gives rise to two alternatively spliced transcripts
A: Reference number: S27391; MUID:93130905; PMID:1483465
A: Accession: S27391
A: Molecule type: DNA
A: Residues: 707-867 <P>
A: Cross-references: EMBL:X68801
A: Note: the authors translated the codon AAA for residue 714 as Leu and AAG for residue 715 as Lys. The EF hand structures are predicted to be incapable of binding calcium.
C: Comment: The EF hand structures are predicted to be incapable of binding calcium.
C: Comment: The amino-terminal domain is involved in actin binding, the middle domain in actin binding.
C: Genetics:
A: Introns: 721/3; 770/3; 792/3; 845/3
A: Note: the list of introns is incomplete
C: Complex: homodimer
C: Superfamily: alpha-actinin; alpha-actinin actin-binding domain homology; calmodulin repeat domain homology
C: Keywords: actin binding; alternative splicing; duplication; EF hand; homodimer; skeleton
F: 40-253/Domain: alpha-actinin actin-binding domain homology <ACT>
F: 283-394/Domain: spectrin/dystrophin repeat homology <SP1>
F: 403-509/Domain: spectrin/dystrophin repeat homology <SP2>
F: 518-630/Domain: spectrin/dystrophin repeat homology <SP3>
F: 639-743/Domain: spectrin/dystrophin repeat homology <SP4>
F: 756-788/Domain: calmodulin repeat homology <EF1>
F: 824-867/Domain: calmodulin repeat homology <EF2>

Query Match 47.28; Score 51; DB 2; Length 897;

Query match	Pred. No. 16;	Gaps
Best Local Similarity 52.6%	0;	0;
Mismatches	7;	Indels
Conservative	2;	Mismatches
10. Conservative	0;	Gaps

2 LPPEMSGLLKRRPQGGP 20
957 LPPEAGQYCIKRMPOVYTG 875

RESULT 3

72716 - Aeropyrum pernix (strain K1)

C:Species: *Aeropyrum pernix* #text change 20-Aug-1999
 hypotnetical protein amino acid change 20-Aug-1999
 #text change 20-Aug-1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain ATCC 79660.
Accession: E72716
Date: 20-Aug-1999
Revision: 20 May 1999

A; Reference number: E72716

A; Status: preliminary

A: Molecule type: DNA
A: Residues: 1-125 <KAW>
A: Cross-references: DDBJ: AP000060; NID: g5104188; PIDN: BAA80133.1; PID: d1043919; PID: g5104188
A: Experimental source: strain K1
C: Genetics:
A: Gene: APE1148

Query Match	46.38; Score 50; DB 2; Length 125;
-------------	------------------------------------

Query matches			
Best Local Similarity	54.5%	Pred. No. 2.9;	
Watched 12: Conservative		2: Mismatches	

Matches 12, conservative 20

QY 1 PLPPMSG--SLKKRPQGEGP 20
||| : | ||| : ||
pb 10 PLLPQDDGRPRLLRRGPQGGGP 31

RESULT 4

ENKVGDBZ

core protein p19 - Cymbidium ringspot virus

C:Species: Cymbidium ringspot virus

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 21-Jul-2000

C:Accession: JS0271; S05459

R:Grieco, F.; Burgvan, J.; Russo, M.

J. Gen. Virol. 70, 2533-2538, 1989

A:Title: Nucleotide sequence of the 3'-terminal region of Cymbidium ringspot virus RN

A:Reference number: A92801; MUID:89381709; PMID:2778443

A:Accession: JS0271

A:Molecule type: genomic RNA

A:Residues: 1-172 <GR1>

R:Grieco, F.; Burgvan, J.; Russo, M.

Nucleic Acids Res. 17, 6383, 1989

A:Title: The nucleotide sequence of Cymbidium ringspot virus RNA.

A:Reference number: JS0268; MUID:89366663; PMID:2771646

A:Accession: S03459

A:Status: preliminary; translation not shown

A:Molecule type: genomic RNA

A:Residues: 1-172 <GR2>

A:Cross-references: EMBL:X15511; NID:g59020; PIDN:CAA33535.1; PID:g4469164

C:Comment: The genome is a single-stranded, positive-sense RNA.

C:Superfamily: tombusvirus core protein p19

C:Keywords: core protein

Query Match	45.48%
Best Local Similarity	47.18%
Matches	8; Conservative
QY	1 PLPPMGSLLKRRPQG 17
Db	147 PTPSREGNVSRRRPEG 163

RESULT 5

EDBE1F
 Immediate-early protein IE180 - suid herpesvirus 1 (strain Indiana-Funkhauser)
 C:Species: suid herpesvirus 1
 C:Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 16-Feb-1997
 C:Accession: S04713
 R:Cheung, A. K.
 Nucleic Acids Res. 17, 4637-4646, 1989
 A:Title: DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies virus
 Accession_number: S04713; MUID:89315207; PMID:2546124

A:Reference number: S04713
A:Accession: S04713
A:Molecule type: DNA
A:Residues: 1-1460 <CHE>
C:Superfamily: herpesvirus immediate-early protein IE175
C:Keywords: DNA binding; early protein; transcription regulation

Query Match	45.4%	Score 49;	DB 1;	Length 1460;
Best Local Similarity	45.0%	Pred. NO. 53;		
Matches	9;	Conservative	2;	Mismatches 9; Indels 0; Gaps 0;

QY	1	PLPPEMSGSLKRRPQGE	20
		:	
	180	PAPPAPPAPRRPRGDG	199

RESULT 6

S37671
 C:Species: Homo sapiens (man)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 15-Sep-2000
 C:Accession: S37671
 R:Bougueret, L. EMBL Data Library, August 1992
 A:Reference number: S37671
 A:Accession: S37671
 A:Status: preliminary
 A:Molecule type: DNA

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: B84430
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-138 <STO>
 A;Cross-references: GB:AE002093; NID:g4522006; PIDN:AAD21779.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: At2g01870
 A;Map position: 2

Query Match 43.1%; Score 46.5; DB 2; Length 138;
 Best Local Similarity 52.6%; Pred. No. 11;
 Matches 10; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 3 PPEMSGSLKRR-RPQGECP 20
 |||:||||| :| :||
 Db 89 PPEISGSFLTWIKPSMDGP 107

RESULT 12

T45392

hypothetical protein [imported] - Mycobacterium leprae

C;Species: Mycobacterium leprae

C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 31-Jan-2000

C;Accession: T45392

R;Cole, S.T.; Flesselles, B.; Honore, N.

submitted to the EMBL Data Library, August 1997

A;Reference number: 222966

A;Accession: T45392

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-572 <COL>

A;Cross-references: EMBL:Z98756; PIDN:CAB11462.1

A;Experimental source: cosmid B2492

C;Genetics:

A;Note: MLCB2492.30

Query Match

Best Local Similarity 43.1%; Score 46.5; DB 2; Length 572;

Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 3 PPEMSGSLKRRPQGECP 20
 |||:||||| :| :||
 Db 10 PPEVHALLSRGP-GSGP 26

RESULT 13

NKVGCU

core protein p20 - cucumber necrosis virus

C;Species: cucumber necrosis virus

A;Note: host Cucumis sativus (cucumber)

C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 23-Jul-1999

C;Accession: JA0133

R;Rochon, D.M.; Tremaine, J.H.

Virology 169, 251-259, 1989

A;Title: Complete nucleotide sequence of the cucumber necrosis virus genome.

A;Reference number: A94391; MUID:89204896; PMID:2705296

A;Accession: JA0133

A;Molecule type: genomic RNA

A;Residues: 1-173 <ROC>

A;Cross-references: GB:M25270; NID:g323338; PIDN:AAA42906.1; PID:g323343

C;Comment: The genome is a single-stranded, positive-sense RNA. It codes for 33K, 92K,

are core proteins.

C;Superfamily: tombusvirus core protein p19

C;Keywords: core protein

Query Match

Best Local Similarity 42.6%; Score 46; DB 1; Length 173;

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 PPEMSGSLKRRPQGECP 17
 |||:||||| :| :|||

Db 149 PCEVEGNVSRRRPQG 163

RESULT 14

A84741

probable myrosinase binding protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C;Accession: A84741

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: A84741

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-471 <STO>

A;Cross-references: GB:AE002093; NID:g2924784; PIDN:AAC04913.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g33070

A;Map position: 2

Query Match

Best Local Similarity 42.6%; Score 46; DB 2; Length 471;

Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRPQGECP 20
 |||:||||| :| :|||
 Db 145 PATPKLRGKWKVEQKGECP 164

RESULT 15

S76563

hypothetical protein - Synecocystis sp. (strain PCC 6803)

C;Species: Synecocystis sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C;Accession: S76563

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis

s.

A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S76563

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-501 <KAN>

A;Cross-references: EMBL:D64002; GB:AB001339; NID:g1001612; PIDN:BAA10409.1; PID:d101

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match

Best Local Similarity 42.6%; Score 46; DB 2; Length 501;

Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRPQGECP 17
 |||:||||| :| :|||
 Db 212 PLPESQGTGTFITRAPOG 228

Search completed: April 8, 2003, 10:52:39

Job time : 5.52983 secs

```

A:Residues: 1-1870 <BOU>
A:Cross-references: EMBL:Z15025; NID:g29374; PID:g29375
C:Genetics:
A:Map position: 6p21.3
A:Introns: 38/2; 97/2; 129/3; 154/1; 202/1; 252/3; 279/2; 327/1; 357/2; 429/3; 588/1; 658/1
C:Superfamily: collagen alpha 1(IV) chain

Query Match 45.4%; Score 49; DB 2; Length 1870;
Best Local Similarity 47.4%; Pred. No. 68;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQGE 20
    ||| | | :|:| |
Db 1571 LPPHSSGFLGSKPGPG 1589

RESULT 7
S36152
MHC class III histocompatibility antigen HLA-B-associated protein 2 [similarity] - human
C:Species: Homo sapiens (man)
C:Date: 06-Jun-1995 #sequence_revision 17-Nov-1995 #text_change 15-Sep-2000
C:Accession: S36152
R:Iris, F.J.M.; Bougueleret, L.; Prieur, S.; Caterina, D.; Primas, G.; Perrot, V.; Jurka
Nature Genet. 3, 137-145, 1993
A:Title: Dense Alu clustering and a potential new member of the NFKappaB family within a
A:Reference number: S36152; MUID:93272029; PMID:8499947
A:Accession: S36152
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1872 <IRI>
A:Cross-references: EMBL:Z15025
A:Note: in the authors' translation residues 32-34 are shown after residue 4 and, conse
A:Note: the authors translated the codon AAT for residue 1000 as His
C:Genetics:
A:Introns: 38/2; 97/2; 129/3; 154/1; 202/1; 252/3; 279/2; 327/1; 357/2; 429/3; 588/1; 658/1
C:Superfamily: collagen alpha 1(IV) chain

Query Match 45.4%; Score 49; DB 2; Length 1872;
Best Local Similarity 47.4%; Pred. No. 68;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQGE 20
    ||| | | :|:| |
Db 1573 LPPHSSGFLGSKPGPG 1591

RESULT 8
B35098
MHC class III histocompatibility antigen HLA-B-associated protein 2 [imported] - human
C:Species: Homo sapiens (man)
C:Date: 10-Aug-1990 #sequence_revision 06-Nov-1992 #text_change 24-Aug-2001
C:Accession: B35098
R:Banerji, J.; Sands, J.; Strominger, J.L.; Spies, T.
Proc. Natl. Acad. Sci. U.S.A. 87, 2374-2378, 1990
A:Title: A gene pair from the human major histocompatibility complex encodes large proli
A:Reference number: A35098; MUID:90192810; PMID:2156268
A:Accession: B35098
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2142 <BAN>
A:Cross-references: GB:M33509; NID:g179338; PIDN:AAA35585.1; PID:g179339; GB:M31293
A:Note: the authors translated the codon AGT for residue 97 as Gly
C:Superfamily: collagen alpha 1(IV) chain

Query Match 45.4%; Score 49; DB 2; Length 2142;
Best Local Similarity 47.4%; Pred. No. 78;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQGE 20
    ||| | | :|:| |
Db 1561 LPPHSSGFLGSKPGPG 1579

A:Residues: 1-147 <MCL>
A:Cross-references: EMBL:Z69240; PIDN:CAA93231.1; GSPDB:GN00066; SPDB:SPAC26A3.08
A:Experimental source: strain 972h; cosmid 26A3
C:Genetics:
A:Gene: SPDB:SPAC26A3.08
A:Map position: 1
A:Introns: 1/3

Query Match 43.5%; Score 47; DB 2; Length 147;
Best Local Similarity 47.8%; Pred. No. 9.7;
Matches 11; Conservative 1; Mismatches 7; Indels 4; Gaps 1;

QY 1 PLPPEMSGSLK-----RRPQGE 19
    | : | | | | | | | | |
Db 85 PMDFSMRGSLLSGPGVARPAGRG 107

RESULT 10
S74712
coproporphyrinogen oxidase (EC 1.3.3.3) III - Synecocystis sp. (strain PCC 6803)
N:Alternate names: coproporphyrinogenase; protein sl11185
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S74712
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocys
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74712
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-340 <KAN>
A:Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAAL6863.1; PID:g165
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: hemF
C:Superfamily: coproporphyrinogen oxidase
C:Keywords: oxidoreductase

Query Match 43.5%; Score 47; DB 2; Length 340;
Best Local Similarity 55.6%; Pred. No. 23;
Matches 10; Conservative 3; Mismatches 1; Indels 4; Gaps 1;

QY 2 LPPEMSGSLKRRPQGE 19
    ||| | | :|:| | | |
Db 87 LPP-----SILKORPEAEG 100

RESULT 11
B84430
hypothetical protein At2g01870 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: B84430
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
```

LENGTH: 461
TYPE: PRT
ORGANISM: Homo sapiens
US-09-894-924-17

Query Match 15.58; Score 349.5; DB 10; Length 461;
Best Local Similarity 29.78; Pred. No. 3.1e-13;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

```
QY 9 APGLAMGFLVIGLFGLLASQPAVP-----PYASE-NOTCRDQEKREYEPQHRICCSR 61
DB 2 APVAVMALAVGL-ELMAAA--HALPAQVAFPTPAPEPGSTCR--LREYDQTAQMCCSK 56
QY 62 CPPGTYSAKCSRIIDYVCATCAENSYNHNHYLFIQOLCR---PCDPVMGLEIAPCTS 118
DB 57 CSPGQHAQVFCITKTSIDYCDSEDESTYQLMNMWPECLSCGRCSQDV---ETQACTR 112
QY 119 KRKTQCRCQPGMFCAMALE-CTHCELLSDCPPG-----TEAEKDEVGKGNHCVPCKA 172
DB 113 EONRICTCRPGMYCALSKQEGCRCLCAPLRKCRPGFVARPGTETSDVY-----CKPCAP 166
QY 173 GHFONTSSPSARCOPHTRCENGIVLEAAGTAQSDTTC--KNPLEPLPEMSGSL----- 225
DB 167 GTFSTSTSDICRPHQICN---VVALPGNASRDVAVCTSTSPTRSKAP---GAVHLPDP 219
QY 226 LKRRPQEGEPNPVAGSWEPKRAH-----PYFPDLVQPLPISGQVS--PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTSTFLLPMGPS--PPAEGSTGDFALPVLIVGVTAIG 272
QY 277 VLEAGVPO-----QOSPLDLTFREPOLE--PGEOSQVAHGNTGIIHV-----TGSMT 320
DB 273 LLITGVVNCVIMTQYKKKPLCLQREAKVPHLPADKARGTQGPQQHLITAPSSSSSSSLE 332
QY 321 ITGNIYIYNGPVLGGPPGPG 340
DB 333 SSASALDRRAPTRNQPAAG 352
```

Search completed: April 8, 2003, 11:03:50
Job time : 45.852 secs

RESULT 14
 US-09-896-096A-17
 : Sequence 17, Application us/09896096a
 : Patent No. US20020061559A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: ASHKENAZI, AVI J
 : APPLICANT: BORSTEIN, DAVID
 : APPLICANT: DODGE, KELLY H.
 : APPLICANT: GURENEY, AUSTIN L.
 : APPLICANT: KIM, KYUNG JIN
 : APPLICANT: LAWRENCE, DAVID A.
 : APPLICANT: PITTI, ROBERT
 : APPLICANT: ROY, MARGARET A
 : APPLICANT: THOMAS, DANIEL B
 : APPLICANT: WOOD, WILLIAM I.
 : TITLE OF INVENTION: DCR3 Polypeptide, A TNFR Homolog
 : FILE REFERENCE: P1134R2 REVISED
 : CURRENT APPLICATION NUMBER: US/09/896,096A
 : CURRENT FILING DATE: 2001-06-28

RESULT 15
US-09-894-924-17
; Sequence 17, Application US/09894924
; Patent No. US20020065210A1

1 APPLICANT: ASHKENAZI, AVI J
 2 APPLICANT: BOTSHEIN, DAVID
 3 APPLICANT: DODGE, KELLY H.
 4 APPLICANT: GURNEY, AUSTIN L.
 5 APPLICANT: KIM, KYUNG JIN
 6 APPLICANT: LAWRENCE, DAVID A.
 7 APPLICANT: PITTI, ROBERT
 8 APPLICANT: ROY, MARGARET A
 9 APPLICANT: THOMAS, DANIEL B
 10 APPLICANT: WOOD, WILLIAM I.
 11
 12 TITLE OF INVENTION: DCR3 POLYPEPTIDE, A TNFR HOMOLOGY
 13
 14 FILE REFERENCE: P1134R2 REVISED
 15 CURRENT APPLICATION NUMBER: US/09/894,924
 16
 17 CURRENT FILING DATE: 2001-06-28
 18
 19 PRIOR APPLICATION NUMBER: US 09/157,289
 20
 21 PRIOR FILING DATE: 1998-09-18
 22
 23 PRIOR APPLICATION NUMBER: US 60/059,288
 24
 25 PRIOR FILING DATE: 1997-09-18
 26
 27 PRIOR APPLICATION NUMBER: US 60/094,640
 28
 29 PRIOR FILING DATE: 1998-07-30
 30
 31 NUMBER OF SEQ ID NOS: 18
 32
 33 SEQ ID NO 17

```

: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BROWDY AND NEIMARK
: STREET: 419 Seventh Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/800,909
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/476,862
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: IL 94039
: FILING DATE: 06-APR-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: IL 91229
: FILING DATE: 06-AUG-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: IL 90339
: FILING DATE: 18-MAY-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: BROWDY, Roger L.
: REGISTRATION NUMBER: 25,618
: REFERENCE/DOCKET NUMBER: WALLACH-12A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-737-3528
: TELEFAX: 202-737-3528
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 461 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-800-909-2

Query Match          15.5%; Score 349.5; DB 10; Length 461;
Best Local Similarity 29.7%; Pred. No. 3.1e-13;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 ABGLAMGPLVLGLFGLAASOPQAVP-----PYASE-NQTCRQDEKEYEPQHRICSR 61
DB 2 AVVAWMAALAVGL-ELMAAA--HALPAQVAFPPVAPPEPSTCR--LREYYDQTQMCCSK 56
QY 62 CPGPTVYSAKCSRINDYVATCAENSYNHNMYLITICQLCR---PCDPYMGLEETAPCTS 118
DB 57 CSPGQHAKEYCTKTSJDYVDCSDSTYTOLMNMWVECLSCGSRCSDDV---ETQACTR 112
QY 119 KRKTCCRCOPGMFCAMALE-CTHCEILSDCPPG-----TEAELKDEYKGNHCVPCA 172
DB 113 EGNRICTCRPGMYCALSKNOGCRILCAPLKKCRPGGVAPRGITSDVY-----CKPCAP 166
QY 173 GHFQNTSSPBARCOPHTRCENGLVEAAGTAOSDTTC--KNPLRLPEMGSGL----- 225
DB 167 GFSTNTSTSDICRPHQICN---VVALPGNMSKDAVCTSTPTSMAP---GAVHLPPQ 219
QY 226 LKRRQGEPPNVASWEPRKAH-----PYFPDLVQPLPLISGVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVLGIVGTALG 272
QY 277 VLEAGVPO-----QOSPLDLTRPOLE--PGEOSOVAHGTNGIHV-----TGSMT 320
DB 273 LLITGVNCVINTQYKKRKLCLQREAKVPHLPADAKAGTQGEQOHLITAPSSSSSILE 332
QY 321 ITGNITYINGPVLGGPPGPG 340

```

```

DB 333 SSASALDRRAPTRNQPOAPG 352

RESULT 12
US-09-826-212-4
: Sequence 4, Application US/09826212
: Patent No. US20010021516A1
: GENERAL INFORMATION:
: APPLICANT: Wei, Ying-Fei
: APPLICANT: Gentz, Reiner
: APPLICANT: Ruben, Steven
: APPLICANT: Ni, Jian
: TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
: FILE REFERENCE: 1488,1280006
: CURRENT APPLICATION NUMBER: US/09/826,212
: CURRENT FILING DATE: 2001-04-05
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: Patentin version 3.0
: SEQ ID NO 4
: LENGTH: 461
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-826-212-4

Query Match          15.5%; Score 349.5; DB 10; Length 461;
Best Local Similarity 29.7%; Pred. No. 3.1e-13;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 ABGLAMGPLVLGLFGLAASOPQAVP-----PYASE-NQTCRQDEKEYEPQHRICSR 61
DB 2 AVVAWMAALAVGL-ELMAAA--HALPAQVAFPPVAPPEPSTCR--LREYYDQTQMCCSK 56
QY 62 CPGPTVYSAKCSRINDYVATCAENSYNHNMYLITICQLCR---PCDPYMGLEETAPCTS 118
DB 57 CSPGQHAKEYCTKTSJDYVDCSDSTYTOLMNMWVECLSCGSRCSDDV---ETQACTR 112
QY 119 KRKTCCRCOPGMFCAMALE-CTHCEILSDCPPG-----TEAELKDEYKGNHCVPCA 172
DB 113 EGNRICTCRPGMYCALSKNOGCRILCAPLKKCRPGGVAPRGITSDVY-----CKPCAP 166
QY 173 GHFQNTSSPBARCOPHTRCENGLVEAAGTAOSDTTC--KNPLRLPEMGSGL----- 225
DB 167 GFSTNTSTSDICRPHQICN---VVALPGNMSKDAVCTSTPTSMAP---GAVHLPPQ 219
QY 226 LKRRQGEPPNVASWEPRKAH-----PYFPDLVQPLPLISGVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVLGIVGTALG 272
QY 277 VLEAGVPO-----QOSPLDLTRPOLE--PGEOSOVAHGTNGIHV-----TGSMT 320
DB 273 LLITGVNCVINTQYKKRKLCLQREAKVPHLPADAKAGTQGEQOHLITAPSSSSSILE 332
QY 321 ITGNITYINGPVLGGPPGPG 340
DB 333 SSASALDRRAPTRNQPOAPG 352

RESULT 13
US-09-758-124-2
: Sequence 2, Application US/09758124
: Patent No. US20020006391A1
: GENERAL INFORMATION:
: APPLICANT: SMITH, Craig A.
: APPLICANT: GOODWIN, Raymond G.
: TITLE OF INVENTION: TUMOR NECROSIS FACTOR-ALPHA AND -BETA RECEPTORS
: FILE REFERENCE: A7895
: CURRENT APPLICATION NUMBER: US/09/758,124
: CURRENT FILING DATE: 2001-01-12
: PRIOR APPLICATION NUMBER: 08/953,268
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 08/555,629

```

NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 1 DNA Way
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/112,793
 FILING DATE: 28-Mar-2002
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/828,683A
 FILING DATE: 31-Mar-1997
 APPLICATION NUMBER: 08/625328
 FILING DATE: 1-Apr-1996
 APPLICATION NUMBER: 08/710802
 FILING DATE: 23-Sep-1996

ATTORNEY/AGENT INFORMATION:
 NAME: Marschang, Diane L.
 REGISTRATION NUMBER: 35,600
 REFERENCE/DOCKET NUMBER: P1007P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650/225-5416
 TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 170 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 14:
 US-10-112-793-14

Query Match 43.9%; Score 987; DB 9; Length 170;
 Best Local Similarity 100.0%; Pred. No. 2e-50;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 42 TCROEKRYEPQHRICCSRCPPGYYSAKSRTIDYTCATCAENSYNEHNNYLTICQLC 101
 |||||
 1 TCROEKRYEPQHRICCSRCPPGYYSAKSRTIDYTCATCAENSYNEHNNYLTICQLC 60
 |||||

QY 102 RPPDVMGLEIETAPCTSKRTQCRQCPGMFCAMALECTHCELLSDCPGTEAEIKDEVG 161
 |||||
 Db 61 RPPDVMGLEIETAPCTSKRTQCRQCPGMFCAMALECTHCELLSDCPGTEAEIKDEVG 120
 |||||

QY 162 KGNHCVCKAGHFQNTSSPSARCPHTRCENOGIYEAPGTAOSDTTC 211
 |||||
 Db 121 KGNHCVCKAGHFQNTSSPSARCPHTRCENOGIYEAPGTAOSDTTC 170
 |||||

RESULT 10
 US-09-768-779A-3
 ; Sequence 3, Application US/09768779A
 ; Patent No. US20020127637A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NI, JIAN
 ; TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR
 ; RECEPTOR-LIKE PROTEIN 8
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
 ; STREET: 9410 KEY WEST AVENUE
 ; CITY: ROCKVILLE
 ; STATE: MD
 ; COUNTRY: US
 ; ZIP: 20850

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/768,779A
 FILING DATE: 25-Jan-2001
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/086,582
 FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: KENLEY K. HOOVER
 REGISTRATION NUMBER: 40,302
 REFERENCE/DOCKET NUMBER: PF368PP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 450 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: Linear

MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-09-768-779A-3

Query Match 15.5%; Score 349.5; DB 10; Length 450;
 Best Local Similarity 29.7%; Pred. No. 3e-13;
 Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APGLAMGPLVLGIFGLIASOPQAVP-----PYASE-NOTCRDQEKRYEPQHRICCS 61
 |||||
 Db 2 APVAVMALAVGL-ELWAAA--HALPAQVAFTPYAPPEPGSTCR--LREYYDOTAKMCCSK 56
 |||||

QY 62 CPPTGYYSAKSRTIDYTCATCAENSYNEHNNYLTICQLC--PCDPVNGLEIETACTS 118
 |||||
 Db 57 CSPQDHAKVFCTKTSIDVCSCEJSTYTQLMNWPPELCSGSSRSSDQV---ETQACTR 112
 |||||

QY 119 KRKQCRQCPGMFCAMALE-CTHCELLSDCPG-----TEAEIKDEVGKGNHCVCKA 172
 |||||
 Db 113 EQNHTICGRPGWCYALSKQECGRICAPLRKCRPGFGVARRPCTERSDVV-----CKPCAP 166
 |||||

QY 173 GHFQNTSSPSARCPHTRCENOGIYEAPGTAOSDTTC--KNLEPLPEMSSGL----- 225
 |||||
 Db 167 GFENHTSSIDICRPHQICN---VVALPGNASRDVACTSTSPTRSMAP--GAVHLPP 219
 |||||

QY 226 LKRPPQGEPPPVAGSMPEPKAH-----PYFPDLVQPLPLPISGVS-PVS--TGLPAAP 276
 |||||
 Db 220 VSTRSQHTQPTP-----EPSTAPSTFLPPGPS--PPABGSTQDFALPVGLIYGVATLG 272
 |||||

QY 277 VLEAGVPQ-----QGSPLDLTREPOLE--PGQSOVAHGTNGIIV-----TGSMT 320
 |||||
 Db 273 ILIGVNCVIMTQVKKKPLCLQREAKVPHLPADKARGTQGPQEOHLLITAPSSSSSSSLE 332
 |||||

QY 321 ITGNIYINGPVLAGPPPG 340
 |||||
 Db 333 SSASALDRAPTRNQPAPG 352
 |||||

RESULT 11
 US-09-800-909-2
 ; Sequence 2, Application US/09800909
 ; Patent No. US20010019833A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WALLACH, David
 ; APPLICANT: BIGDA, Jacek
 ; APPLICANT: BELESKY, Igor
 ; APPLICANT: METT, Igor
 ; APPLICANT: ENGELMANN, Hartmut
 ; TITLE OF INVENTION: TNF INHIBITORS

;; PRIOR APPLICATION NUMBER: 60/121,774
;; PRIOR FILING DATE: 1999-03-04
;; PRIOR APPLICATION NUMBER: 09/006,352
;; PRIOR FILING DATE: 1998-01-13
;; PRIOR APPLICATION NUMBER: 60/035,496
;; PRIOR FILING DATE: 1997-01-14
;; NUMBER OF SEQ ID NOS: 42
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO: 8
;; LENGTH: 415
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-935-727-8

Query Match 62.5%; Score 1404; DB 10; Length 415;
Best Local Similarity 63.9%; Pred. No. 4.3e-74;
Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

QY 1 MLPMATAPGLAMGPLVLGLGLLAASQAPVPPASENOTCRDOEKEYEPOHRIICS 60
DB 1 MRLPRASPCGLAMGPLLGLGLVVASQPOLVPPYRIENGTQMDQREYEPMDVCS 60
QY 61 RCPPTIYSAKSRIRDTVCATCAENSYNEHNNITTCOLCRPCDPVAGLEIAPCTSKR 120
DB 61 RCPPEEFVAVCSRSQDVCCTCPHNSYNEHNNHLSCTOLCRPCDIYVGFEEVAPCTSDR 120
QY 121 KTQCSQGMFCAMALECTHC--ELSDCPPTGAEELKDEYKGNHCVCKAGHPONT 178
DB 121 KAECCQGMSCYIYDNECVHCEERLVLCQPTAEVTDLMQDVAVCKPCKHFOHT 180
QY 179 SSPSARCCPHRTCENOGLVEAAPTASQSDTCKNLEP----- 216
DB 181 SSPRARCOPHRTCEIOGLVEAAPGTSDTCKNPEPGAMLLALLSLVLELFTTVL 240
QY 217 -----LPEMS---GSLKRRPGEGRPVAGSWEPKAPHPYDVLQPLDISGVSPV 268
DB 241 ACAMMRHPSLCKLSTLKRHEGESPPCPA---PRADPEPLAPLPLMSGDLSPS 296
QY 269 STGLPAFAVLEAGVPOQSPDLITREPOLERPEQSOVAHGTCIHVTGSGMTITGNIYI 328
DB 297 PAGPPTASLEEVVLQOQSPVLQARELEAPREHGVAGANGIHVTGSGSVYVNTNIYI 356
QY 329 NGPVIGPPGPDLPATPEPPYPIPEEGDPGPGISTPHQEDGKAMHLETHCG 383
DB 357 NGPVIGTRGPDPPAPPEPPYPTPEEGAPGPELSLTYQEDGKAMHLETHG 411

RESULT 7
US-09-948-018-19
;; Sequence 19, Application US/09948018
;; Patent No. US20020150977A1
;; GENERAL INFORMATION:
;; APPLICANT: Thell et al
;; TITLE OF INVENTION: TNF RECEPTOR-LIKE MOLECULES AND USES THEREOF
;; FILE REFERENCE: 0101737677
;; CURRENT APPLICATION NUMBER: US/09/948, 018
;; CURRENT FILING DATE: 2001-09-05
;; PRIOR APPLICATION NUMBER: US 60/230,191
;; PRIOR FILING DATE: 2000-09-05
;; NUMBER OF SEQ ID NOS: 45
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 19
;; LENGTH: 257
;; TYPE: PRF
;; ORGANISM: Mus musculus
US-09-948-018-19

Query Match 53.9%; Score 1211; DB 10; Length 257;
Best Local Similarity 86.0%; Pred. No. 3.6e-63;
Matches 221; Conservative 0; Mismatches 0; Indels 36; Gaps 1;
QY 90 EHWNTLTICQLCRPCDPVAGLEIAPCTSKRRTQCRQGMFCAMALECTHCELLSDCP 149
|||||

DB 1 EHWNTLTICQLCRPCDPVAGLEIAPCTSKRRTQCRQGMFCAMALECTHCELLSDCP 60
QY 150 PGTEAEIKDEYKGNHCVCKAGHPONTSSPSARCOPTHCENGLVEAAPGTASQDPT 209
DB 61 PGTEAEIKDEYKGNHCVCKAGHPONTSSPSARCOPTHCENGLVEAAPGTASQDPT 120
QY 210 CKNPLEPLPEMS-----GSLKRRPQGE 233
DB 121 CKNPLEPLPEMSGTMMLAVLLPLAFLLATVFSCTIKSHPSLCKRLGSLKRRPQGE 180
QY 234 GPNPVAGSWEPKAPHPYDVLQPLISGVSPVSTGLPAAPVLEAGVPOQSPDLITR 293
DB 181 GPNPVAGSWEPKAPHPYDVLQPLISGVSPVSTGLPAAPVLEAGVPOQSPDLITR 240
QY 294 EPOLERQSOVAHGTCN 310
DB 241 EPOLERQSOVAHGTCN 257

RESULT 8
US-10-003-211-1
;; Sequence 1, Application US/10003211
;; Publication No. US20020197254A1
;; GENERAL INFORMATION:
;; APPLICANT: Biogen, Inc.
;; TITLE OF INVENTION: Soluble Lymphotoxin Beta Receptor and
;; TITLE OF INVENTION: Anti-Lymphotoxin Receptor and Ligand Antibodies as
;; TITLE OF INVENTION: Therapeutic Agents for the treatment of Immunological
;; TITLE OF INVENTION: Diseases
;; FILE REFERENCE: A01305
;; CURRENT APPLICATION NUMBER: US/10/003, 211
;; CURRENT FILING DATE: 2001-10-31
;; PRIOR APPLICATION NUMBER: PCT/US97/19436
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/029, 060
;; PRIOR FILING DATE: 1996-10-25
;; NUMBER OF SEQ ID NOS: 1
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1
;; LENGTH: 197
;; TYPE: PRF
;; ORGANISM: Homo Sapien
US-10-003-211-1

Query Match 50.2%; Score 1129; DB 9; Length 197;
Best Local Similarity 99.5%; Pred. No. 1.5e-58;
Matches 196; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 28 SQPQAVPPYASENOTCRDOEKEYEPOHRIICSRCPGTIYSAKSRIRDTVCATCAENS 87
DB 1 SQPQAVPPYASNOTCRDOEKEYEPOHRIICSRCPGTIYSAKSRIRDTVCATCAENS 60
QY 88 YNEHNNVLTICQLCRPCDPVAGLEIAPCTSKRRTQCRQGMFCAMALECTHCELLSD 147
DB 61 YNEHNNVLTICQLCRPCDPVAGLEIAPCTSKRRTQCRQGMFCAMALECTHCELLSD 120
QY 148 CPGTEAEIKDEYKGNHCVCKAGHPONTSSPSARCOPTHCENGLVEAAPGTASQD 207
DB 121 CPGTEAEIKDEYKGNHCVCKAGHPONTSSPSARCOPTHCENGLVEAAPGTASQD 180
QY 208 TTCKNPLEPLPEMSGS 224
DB 181 TTCKNPLEPLPEMSGT 197

RESULT 9
US-10-112-793-14
;; Sequence 14, Application US/10112793
;; Publication No. US20020192729A1
;; GENERAL INFORMATION:
;; APPLICANT: Ashkenazi, Avi J.
;; TITLE OF INVENTION: Apo-2 LI AND Apo-3 POLYPEPTIDES

QY 325 IYINCPVLGGPPGDDLPATPEPPYPIPEEGDPGPGISTPHQEDGKAMHLAETHECGA 384
 |||||
 Db 361 IYINCPVLGGPPGDDLPATPEPPYPIPEEGDPGPGISTPHQEDGKAMHLAETHECGA 420
 |||||
 QY 385 TFSNRGRNPFTHD 399
 |||||
 Db 421 TFSNRGRNPFTHD 435

RESULT 4

US-09-826-212-6
 ; Sequence 6, Application US/09826212
 ; Patent No. US20010021516A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wei, Ying-Pei
 ; APPLICANT: Gentz, Reiner
 ; APPLICANT: Ruben, Steven
 ; APPLICANT: Ni, Jian
 ; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
 ; FILE REFERENCE: 1488,1280006
 ; CURRENT APPLICATION NUMBER: US/09/826,212
 ; CURRENT FILING DATE: 2001-04-05
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 415
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-826-212-6

Query Match 62.5%; Score 1404; DB 10; Length 415;
 Best Local Similarity 63.9%; Pred. No. 4,3e-74;

Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

QY 1 MLPMWTSAPGLAMGVLVIGLGLAASQPAVPYASENQTCRDOKEKEYEPQHRTCCS 60
 |||||
 Db 1 MRLPRASSPCGLAMGVLVIGLGLAASQPAVPYRIENQTCWDQKEYEPEMHVDCS 60
 |||||
 QY 61 RCPGGTYVSAKCSIRIDTVCATCAENSYNEHMYLTICQLCRPDVPMGLEIAPCTSKR 120
 |||||
 Db 61 RCPGGEVEFAVCSRSODTYCKTCPHNSYNEMHNLSTCOLCRPCDYLVEFEVAPCTSDR 120
 |||||
 QY 121 KTCRCQPGMFCAMALECTHC--ELLSDCPGTEAELEKDEYKGNHNCVCKAGHFOHT 178
 |||||
 Db 121 KACRCQPGMFCAMALECTHC--ELLSDCPGTEAELEKDEYKGNHNCVCKAGHFOHT 180
 |||||
 QY 179 SSPSARCQPHTRCENQGLVEAPGTAQSDTTCNPLEP----- 216
 |||||
 Db 181 SSPSARCQPHTRCENQGLVEAPGTAQSDTTCNPLEP----- 240
 |||||
 QY 217 -----LPEPMS--GSLTKRRPQGEGRPPVAGSMPEPKAHYFPDVLVOPLLPISGDVSPV 268
 |||||
 Db 241 ACAMMRHPSLCRKLGLTKRHPGEESPPCPA---PRADHPFDLAEPLLPMSGDLSPS 296
 |||||
 QY 269 STGLPAAPVLEAGVPOQSPDLITREPQLEPGESQVAGHNGIHTVGGSMITGNIIYI 328
 |||||
 Db 297 PAGPPTAPSLAEVYLQQSPVLQVARELEAEPGEGHGVANGIHVGGSVTVGNIIYI 356
 |||||
 QY 329 NGPVLGGPPGDDLPATPEPPYPIPEEGDPGPGISTPHQEDGKAMHLAETHECG 383
 |||||
 Db 357 NGPVLGGPPGDDLPATPEPPYPIPEEGDPGPGISTPHQEDGKAMHLAETHECG 411

RESULT 5

US-09-907-372-20
 ; Sequence 20, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti G.
 ; APPLICANT: Warren, Bridget A.
 ; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
 ; FILE REFERENCE: PC-0050 US
 ; CURRENT APPLICATION NUMBER: US/09/907,372

;; CURRENT FILING DATE: 2001-07-27
 ;; NUMBER OF SEQ ID NOS: 20
 ;; SOFTWARE: PERL Program
 ;; SEQ ID NO 20
 ;; LENGTH: 415
 ;; TYPE: PRT
 ;; ORGANISM: Mus musculus
 ;; FEATURE:
 ;; NAME/KEY: misc feature
 ;; OTHER INFORMATION: Incyte ID NO. US20020068242A1 9600223
 ; US-09-907-372-20

Query Match 62.5%; Score 1404; DB 10; Length 415;
 Best Local Similarity 63.9%; Pred. No. 4,3e-74;

Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

QY 1 MLPMWTSAPGLAMGVLVIGLGLAASQPAVPYASENQTCRDOKEKEYEPQHRTCCS 60
 |||||
 Db 1 MRLPRASSPCGLAMGVLVIGLGLAASQPAVPYRIENQTCWDQKEYEPEMHVDCS 60
 |||||
 QY 61 RCPGGTYVSAKCSIRIDTVCATCAENSYNEHMYLTICQLCRPDVPMGLEIAPCTSKR 120
 |||||
 Db 61 RCPGGEVEFAVCSRSODTYCKTCPHNSYNEMHNLSTCOLCRPCDYLVEFEVAPCTSDR 120
 |||||
 QY 121 KTCRCQPGMFCAMALECTHC--ELLSDCPGTEAELEKDEYKGNHNCVCKAGHFOHT 178
 |||||
 Db 121 KACRCQPGMFCAMALECTHC--ELLSDCPGTEAELEKDEYKGNHNCVCKAGHFOHT 180
 |||||
 QY 179 SSPSARCQPHTRCENQGLVEAPGTAQSDTTCNPLEP----- 216
 |||||
 Db 181 SSPSARCQPHTRCENQGLVEAPGTAQSDTTCNPLEP----- 240
 |||||
 QY 217 -----LPEPMS--GSLTKRRPQGEGRPPVAGSMPEPKAHYFPDVLVOPLLPISGDVSPV 268
 |||||
 Db 241 ACAMMRHPSLCRKLGLTKRHPGEESPPCPA---PRADHPFDLAEPLLPMSGDLSPS 296
 |||||
 QY 269 STGLPAAPVLEAGVPOQSPDLITREPQLEPGESQVAGHNGIHTVGGSMITGNIIYI 328
 |||||
 Db 297 PAGPPTAPSLAEVYLQQSPVLQVARELEAEPGEGHGVANGIHVGGSVTVGNIIYI 356
 |||||
 QY 329 NGPVLGGPPGDDLPATPEPPYPIPEEGDPGPGISTPHQEDGKAMHLAETHECG 383
 |||||
 Db 357 NGPVLGGPPGDDLPATPEPPYPIPEEGDPGPGISTPHQEDGKAMHLAETHECG 411

RESULT 6

US-09-935-727-8
 ; Sequence 8, Application US/09935727
 ; Patent No. US20020150583A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Human Genome Sciences, Inc.
 ; TITLE OF INVENTION: Tumor Necrosis Factor Receptors 6 Alpha and 6 Beta
 ; FILE REFERENCE: PR454P2
 ; CURRENT APPLICATION NUMBER: US/09/935,727
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: 60/303,224
 ; PRIOR FILING DATE: 2001-07-06
 ; PRIOR APPLICATION NUMBER: 60/252,131
 ; PRIOR FILING DATE: 2000-11-21
 ; PRIOR APPLICATION NUMBER: 60/227,598
 ; PRIOR FILING DATE: 2000-08-25
 ; PRIOR APPLICATION NUMBER: 09/518,931
 ; PRIOR FILING DATE: 2000-03-03
 ; PRIOR APPLICATION NUMBER: 60/168,235
 ; PRIOR FILING DATE: 1999-12-01
 ; PRIOR APPLICATION NUMBER: 60/146,371
 ; PRIOR FILING DATE: 1999-08-02
 ; PRIOR APPLICATION NUMBER: 60/131,964
 ; PRIOR FILING DATE: 1999-04-30
 ; PRIOR APPLICATION NUMBER: 60/131,270
 ; PRIOR FILING DATE: 1999-04-27
 ; PRIOR APPLICATION NUMBER: 60/124,092
 ; PRIOR FILING DATE: 1999-03-12

QY	241	SWEPPKAAHPERDLYOPLPLPSGVSPVSYNGLPAAVYLEAGVGQQOSPLDLTREPOLEPG	300
Db	241	SWEPPKAAHPERDLYOPLPLPSGVSPVSYNGLPAAVYLEAGVGQQOSPLDLTREPOLEPG	3000
QY	301	EOSVAVGCTNGIHTVGGSMITGNIITYNNGPVLCGGPGPGLDLPATPEPPVPIREEGDPCP	360
Db	301	EOSVAVGCTNGIHTVGGSMITGNIITYNNGPVLCGGPGPGLDLPATPEPPVPIREEGDPCP	3600
QY	361	PGISTPHEQEDGKAHWLAETEHCGATPSNRGRPRNOFITHD	399
Db	361	PGISTPHEQEDGKAHWLAETEHCGATPSNRGRPRNOFITHD	399

RESULT 2

```

US-09-907372-19
Sequence 19, Application US/09907372
Patent No. US20020068242A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti G.
TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
FILE REFERENCE: PC-0050 US
CURRENT APPLICATION NUMBER: US/09/907,372
CURRENT FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 19
LENGTH: 435
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020068242A1_g339762
US-09-907-372-19

```

Query Match	98.8%	Score 2230	DB 10	Length 435
Best Local Similarity	91.7%	Pred. No. 3.9e-121		
Matches 399	Conservative 0	Mismatches 0	Indels 36	Gaps 1

QY	1	MLLWMAISAPILANGPVLVLGFLGFLAASQPAQVAPVYASENOTCSDRDEKEYEFOHRI CCS	60
Db	1	MLLWMAISAPILANGPVLVLGFLGFLAASQPAQVAPVYASENOTCSDRDEKEYEFOHRI CCS	60
QY	61	RCPPGTIVYSAKCSIRIDTVCATCAENSYNEHWNLYITCOLCRPCDPVYMLLEETAPCTSKR	120
Db	61	RCPPGTIVYSAKCSIRIDTVCATCAENSYNEHWNLYITCOLCRPCDPVYMLLEETAPCTSKR	120
QY	121	KTCQRCQPGMCAAMALECHTCELLSQCPRTEALEKDEYKGNHHCVPCKAGHFQNTSS	180
Db	121	KTCQRCQPGMCAAMALECHTCELLSQCPRTEALEKDEYKGNHHCVPCKAGHFQNTSS	180
QY	181	PSARCQPHTRRENGGLVEAARAGTAAQSDTCKNPLPELPIRPEMS-----2222	
Db	181	PSARCQPHTRRENGGLVEAARAGTAAQSDTCKNPLPELPIRPEMSGTMMLAVLLPLAFLL	240
QY	223	-----GSLKRRPQGEQGNPVYASGMEPPKAPHYRDDLVLPLISGD	264
Db	241	ATVESCILKSHPSLCKRLGSLILKRRPDQEGENPVYASGMEPPKAPHYRDDLVLPLISGD	300
QY	265	VSPYSTGIPAPVYLEAGVPOQOOSPLDLTRFEQLERPGQSOVAHGTNGIHTVGSMTTGN	324
Db	301	VSPYSTGIPAPVYLEAGVPOQOOSPLDLTRFEQLERPGQSOVAHGTNGIHTVGSMTTGN	360
QY	325	IYIYNGPVLGSPGPGDLPATPEPPYPPIPEBGDGPBGGLSTPHQEDCKAMHLAETHEGCA	384
Db	361	IYIYNGPVLGSPGPGDLPATPEPPYPPIPEBGDGPBGGLSTPHQEDCKAMHLAETHEGCA	420
QY	385	TPSNRGPANOFITHD	399
Db	421	TPSNRGPANOFITHD	435

RESULT 3

US-09-768-779A-6
Sequence 6, Application US/09768779A
Patent No. US20020127637A1
GENERAL INFORMATION:
APPLICANT: NI, JIAN
MOORE, PAUL
TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR
RECEPTOR-LIKE PROTEIN 8
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/768,779A
FILING DATE: 25-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/086,582
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: KENLEY K. HOOVER
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PF368BP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-768-779A-6

Query Match	98.8%	Score 2220:	DB 10:	Length 435:
Best Local Similarity	91.7%:	Pred. No. 3.9e-121:		
Matches 399:	Conservative	0:	Mismatches 0:	Indels 36:
				Gaps 1:
Qy	1	MLLPATSAFGLAMQPLVYLGLFGLLAASQPAQVAPVYASENOTCRDQEKYEYEPQRIKCS	60	
Db	1	MLLPATSAFGLAMQPLVYLGLFGLLAASQPAQVAPVYASNOTCRQEKYEYEPQRIKCS	60	
Qy	61	RCPPGTYSAKCSRIRDYVCATCAENSYNEHWNVLTICOLRCPCDPMGLIEIAPCTSKR	120	
Db	61	RCPPGTYSAKCSRIRDYVCATCAENSYNEHWNVLTICOLRCPCDPMGLIEIAPCTSKR	120	
Qy	121	KTCQCGQGMFCANAALECTHCELLSDCPGTCGEALKEIDYVGKGNHCHVCKRGHHQNTSS	180	
Db	121	KTCQCGQGMFCANAALECTHCELLSDCPGTCGEALKEIDYVGKGNHCHVCKRGHHQNTSS	180	
Qy	181	PSARCQPHTRCENOGILVEAPGTAOSDTCTCKNPLEPLREPMSS-----	222	
Db	181	PSARCQPHTRCENOGILVEAPGTAOSDTCTCKNPLEPLREPMSSGTMILMLAVLLPLAFLLL	240	
Qy	223	-----GSLIKRRPQEGGPNVAGSWEPKRAHFYFPDLVQPLPLPISGD	264	
Db	241	ATVFCISIMKSHPSLCRKILGSLILKRRPQEGGPNVAGSWEPKRAHFYFPDLVQPLPLPISGD	300	
Qy	265	VSPVSTGTPAAVYELAGVPOQSSPLDLTRREPDLREGESQVAHNGINGIHVYGGSWITTCGN	324	
Db	301	VSPVSTGTPAAVYELAGVPOQSSPLDLTRREPDLREGESQVAHNGINGIHVYGGSWITTCGN	360	

2

2

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:51:46 ; Search time 42.852 Seconds
(without alignments)
569.244 Million cell updates/sec

Title: US-09-917-372-1

Perfect score: 4250
Sequence: 1 MLIPWATSAPGLAWGPLVLG.....EHGATPSNRGRNQFIITHD 399

Scoring table: BLOSUM62
Gapop 10.0 : Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

```

1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.dep:*
2: /cgn2_6/ptodata/1/pubppa/PCr_NEW_PUB.dep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.dep:*
4: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.dep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.dep:*
6: /cgn2_6/ptodata/1/pubppa/PCrUS_PUBCOMB.dep:*
7: /cgn2_6/ptodata/1/pubppa/PCrUS_PUBCOMB.dep:*
8: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.dep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.dep:*
10: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.dep:*
11: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.dep:*
12: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.dep:*
13: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.dep:*
14: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.dep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2248	100.0	399	10	US-09-907-372-1	Sequence 1, Appli
2	2220	98.8	435	10	US-09-907-372-19	Sequence 19, Appli
3	2220	98.8	435	10	US-09-768-779A-6	Sequence 6, Appli
4	1404	62.5	415	10	US-09-826-212-6	Sequence 6, Appli
5	1404	62.5	415	10	US-09-907-372-20	Sequence 20, Appli
6	1404	62.5	415	10	US-09-935-727-8	Sequence 8, Appli
7	1211	53.9	257	10	US-09-948-018-19	Sequence 19, Appli
8	1129	50.2	197	10	US-10-003-211-1	Sequence 1, Appli
9	987	43.9	170	9	US-10-112-793-14	Sequence 14, Appli
10	349.5	15.5	450	10	US-09-768-779A-3	Sequence 3, Appli
11	349.5	15.5	461	10	US-09-800-909-2	Sequence 2, Appli
12	349.5	15.5	461	10	US-09-826-212-4	Sequence 4, Appli
13	349.5	15.5	461	10	US-09-758-124-2	Sequence 2, Appli
14	349.5	15.5	461	10	US-09-896-096A-17	Sequence 17, Appli
15	349.5	15.5	461	10	US-09-894-924-17	Sequence 17, Appli
16	349.5	15.5	461	10	US-09-840-707A-17	Sequence 17, Appli
17	349.5	15.5	461	10	US-09-800-908-3	Sequence 3, Appli
18	349.5	15.5	461	10	US-09-935-727-6	Sequence 6, Appli
19	349.5	15.5	461	12	US-10-164-592-3	Sequence 3, Appli

20	327.5	14.6	474	10	US-09-758-182-4	Sequence 4, Appl 1
21	321.5	14.3	225	10	US-09-840-795-10	Sequence 10, Appl 1
22	319	14.2	300	9	US-10-129-709-3	Sequence 3, Appl 1
23	319	14.2	300	10	US-09-896-096A-1	Sequence 1, Appl 1
24	319	14.2	300	10	US-09-894-924-1	Sequence 1, Appl 1
25	319	14.2	300	10	US-09-935-727-2	Sequence 2, Appl 1
26	314.5	14.0	277	10	US-09-826-212-10	Sequence 10, Appl 1
27	314.5	14.0	277	10	US-09-768-779A-4	Sequence 4, Appl 1
28	314.5	14.0	277	10	US-09-935-727-12	Sequence 12, Appl 1
29	313.5	13.9	277	10	US-09-855-528-2	Sequence 2, Appl 1
30	312	13.9	625	9	US-09-877-650-15	Sequence 15, Appl 1
31	312	13.9	625	9	US-10-166-323A-2	Sequence 2, Appl 1
32	312	13.9	625	10	US-09-871-856-15	Sequence 15, Appl 1
33	311.5	13.9	277	10	US-09-839-339A-1	Sequence 1, Appl 1
34	310.5	13.8	299	10	US-09-877-156-17	Sequence 17, Appl 1
35	310	13.8	300	10	US-09-840-195-2	Sequence 2, Appl 1
36	309.5	13.8	235	9	US-10-243-220-6	Sequence 2, Appl 1
37	309.5	13.8	235	9	US-10-243-220-6	Sequence 8, Appl 1
38	309.5	13.8	235	9	US-09-882-735-16	Sequence 16, Appl 1
39	309.5	13.8	235	10	US-09-907-263-4	Sequence 4, Appl 1
40	308.5	13.7	235	9	US-10-243-220-6	Sequence 4, Appl 1
41	308.5	13.7	235	9	US-10-243-220-6	Sequence 6, Appl 1
42	306	13.6	227	10	US-09-840-795-9	Sequence 9, Appl 1
43	303.5	13.5	271	9	US-10-129-709-1	Sequence 1, Appl 1
44	303	13.5	184	9	US-09-855-455-8	Sequence 8, Appl 1
45	303	13.5	211	10	US-09-877-156-20	Sequence 20, Appl 1

ALIGNMENTS

```

RESULT 1
US-09-907-372-1
; Sequence 1, Application US/09907372
; Patent No. US20020068242A1
;
GENERAL INFORMATION:
;
APPLICANT: Lal, Preeti G.
APPLICANT: Warren, Bridget A.
TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
FILE REFERENCE: PC-0050 US
CURRENT APPLICATION NUMBER: US/09/907,372
CURRENT FILING DATE: 2001-07-27
NUMBER OF SEQ. ID NOS: 20
;
SOFTWARE: PERL Program
;
SEQ ID NO 1
;
LENGTH: 399
;
TYPE: PRT
;
ORGANISM: Homo sapiens
;
FEATURE:
;
NAME/KEY: misc-feature
;
OTHER INFORMATION: Incyte ID No. US20020068242A1 7497867CDD
US-09-907-372-1

```

Query Match	Similarity	100.0%	Score	2248;	DB	103;	Length	399;
Best Local	Similarity	100.0%	Pred.	No. 8.9e-123;				
Matches	399;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps
QY	1	MLLPATSAAGLAMPGLVLTGLFGLLAASQPAVPPYASNQCRCRDEKRYEPQRHTCS	60					
Db	1	MLLPATSAAPGLAMPGLVLTGLFGLLAASQPAQVAPPYASBNQCRCRDEKRYEPQRHTCS	60					
QY	61	RCPPGYVSAKCSRIKDYCATCATCAENSYEHNNVLTICQLCPDCPDVMDLEETAPCTSKR	120					
Db	61	RCPPGYVSAKCSRIKDYCATCATCAENSYEHNNVLTICQLCRCPDPMGLLEETAPCTSKR	120					
QY	121	KTCQRCQPMFCAMALLECTHCELLSDCPPTGEALKEKDEYGVGNHNCVCKKASHPNTSS	180					
Db	121	KTCQRCQPMFCAMALLECTHCELLSDCPPTGEALKEKDEYGVGNHNCVCPKASHPNTSS	180					
QY	181	PSARCPQPHRCENOGIYEAPAPGAOSDTCTCKNPLEPLPEMSSGLLKRDPQGGPMPVAG	240					
Db	181	PSARCPQPHRCENOGIYEAPAPGAOSDTCTCKNPLEPLPEMSSGLLKRDPQGGPMPVAG	240					

Matches 69; Conservative 22; Mismatches 83; Indels 60; Gaps 11;

QY 48 KEYTEPOHRIICSRCPETVYSARCSIRIDVTACATCAENSYNHEMNYLTITQLCRPDDPV 107
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 29 KHTVPSGHK -CCRECOPGHGMYSCRDHRTDLCHPCETGFENAVNNDT-CKCTQCNRH 86
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 108 MGLEIAPCTSKRKTQCRCQCGMFCAAMALECTHCCELLSDCPRGTAEALNKDEVKGNHC 167
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 87 SGSEIKONCTPTDGVTSACR-----PGTPGR-QDSGYRLGDVC 123
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 168 VPCKAGHPONTSSP-SARKCPHTRCENOGLVEAAPGAOSDTTCKNLIEPLPEMGSSL 225
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 124 VPCPPGGH-----SGNNACKCFWTNCILSGKHSPADSIDANVED-----KSLLATL 173
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY 226 LKRKPGEPPNPVAGSWEPRAHPYE-PDVLPPLPIISGDVSPYSTGLPAAPVL 278
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 174 I-----METQR--PTFRPTVO-----STVMPTSLSLPSPPTL 205
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 13

GQHDTI
tumor necrosis factor receptor 1 precursor [validated] - human
N:Alternate names: p55 tumor necrosis factor receptor; TNF receptor type 1
N:Contains: tumor necrosis factor alpha inhibitor; tumor necrosis factor binding protein
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1992 #sequence-revision 30-Jun-1992 #text-change 08-Dec-2000
A:Accession: A38208; A34899; A34900; A36555; C36555; A38281; S12057; J07058; A60231; A38
R:Fuchs, P.; Strehl, S.; Dworkzak, M.; Himmler, A.; Ambros, P.F.
Genomics 13, 219-224, 1992
A>Title: Structure of the human TNF receptor 1 (p60) gene (TNRF1) and localization to ch
A:Reference number: A38208; MUID:92250045; PMID:1315717
A:Accession: A38208
A:Molecule type: DNA
A:Residues: 1-455 <FUOC>
A:Cross-references: GB:M75864; GB:M75865; GB:M75866; NID:G339748; PIDN:AAA61201.1; PID:g
R:Lotscher, H.; Pan, Y.C.E.; Lahm, H.W.; Gentz, R.; Brockhaus, M.; Tschuch, H.; Lesslau
Cell 61, 351-359, 1990
A>Title: Molecular cloning and expression of the human 55 kd tumor necrosis factor recep
A:Reference number: A34899; MUID:90235284; PMID:2158862
A:Accession: A34899
A:Molecule type: mRNA
A:Residues: 1-455 <LOE>
A:Cross-references: GB:M58286; GB:M33480; NID:G339753; PIDN:AAA36753.1; PID:G339754
A:Experimental source: Placenta
A>Note: part of this sequence, including the amino end of the mature protein, confirmed
R:Schall, T.J.; Lewis, M.; Koller, K.J.; Lee, A.; Rice, G.C.; Wong, G.H.W.; Galanaga, T
Cell 61, 361-370, 1990
A>Title: Molecular cloning and expression of a receptor for human tumor necrosis factor
A:Reference number: A34900; MUID:90235285; PMID:2158863
A:Accession: A34900
A:Molecule type: mRNA
A:Residues: 1-455 <SCH>
A:Cross-references: GB:M33294; NID:G339744; PIDN:AAA03210.1; PID:G339745
R:Himmler, A.; Maurer-Fogy, I.; Kroenke, M.; Scheutich, P.; Pfizenmaier, K.; Lantz, M.;
DNA Cell Biol. 9, 705-715, 1990
A>Title: Molecular cloning and expression of human and rat tumor necrosis factor recepted
A:Reference number: A36555; MUID:91090841; PMID:1702293
A:Accession: A36555
A:Molecule type: mRNA
A:Residues: 1-455 <HM>
A:Cross-references: GB:M63121; NID:G339755; PIDN:AAA36754.1; PID:G339756
A:Accession: C36555
A:Molecule type: Protein
A:Residues: 30-38;41-53,'X','55'-79,'XX',82-94,'NK','XX',100-104;107-128;162-167,'X',169-2;
A>Note: The purified protein, called tumor necrosis factor binding protein, is a soluble
R:Gray, P.M.; Barrett, K.; Chantry, D.; Turner, M.; Feldmann, M.
Proc. Natl. Acad. Sci. U.S.A. 87, 7380-7384, 1990
A>Title: Cloning of human tumor necrosis factor (TNF) receptor cDNA and expression of rec
A:Reference number: A38281; MUID:91017509; PMID:2170974
A:Accession: A38281
A:Molecule type: mRNA
A:Residues: 1-455 <GRA>
A:Cross-references: GB:M37764

A:Note: the authors translated the codon TGG for residue 371 as Thr, AAG for residue
 R:Ophnar, Y.; Kemper, O.; Brakebusch, C.; Engelmann, H.; Zwaag, R.; Aderka, D.; Holtu
 EMBO J. 9, 3269-3278, 1990
 A:Title: Soluble forms of tumor necrosis factor receptors (TNF-Rs). The cDNA for the
 le form of the receptor.
 A:Reference number: S12057; MUID:91006021; PMID:1698610
 A:Accession: S12057
 A:Molecule type: mRNA
 A:Residues: 1-455 <NDP>
 A:Cross-references: EMBL:X55313; NID:937223; PIDD:CA939021.1; PID:937224
 A:Note: parts of soluble TNF binding protein 1, including its amino and carboxyl ends
 R:Kemper, O.; Wallach, D.
 Gene 134, 209-216, 1993
 A:Title: Cloning and partial characterization of the promoter for the human p55 tumor
 A:Reference number: J10758; MUID:94085779; PMID:8262379
 A:Accession: J10758
 A:Molecule type: DNA
 A:Residues: 1-13 <KEM>
 R:Seckinger, P.; Vey, E.; Turcatti, G.; Wingfield, P.; Dayer, J.M.
 Eur. J. Immunol. 20, 1167-1174, 1990
 A:Title: Tumor necrosis factor inhibitor: purification, NH-2-terminal amino acid sequ
 A:Reference number: A60231; MUID:90292116; PMID:2113477
 A:Accession: A60231
 A:Molecule type: protein
 A:Residues: 41-43, 'X', 45-53, 'X', 55-57 <SPC>
 R:Galanaga, T.; Hwang, C.; Koh, W.; Cappuccini, F.; Lucel III, J.A.; Jeffes, E.W.B.,
 Proc. Natl. Acad. Sci. U.S.A. 87, 8781-8784, 1990
 A:Title: Purification and characterization of an inhibitor (soluble tumor necrosis f
 A:Reference number: A38258; MUID:91062364; PMID:2174164
 A:Accession: A38258
 A:Molecule type: protein
 A:Residues: 41-60 <GAT>
 A:Experimental source: cancer patient serum
 R:Olsson, I.; Lantze, M.; Nilsson, E.; Peetre, C.; Thyrell, H.; Grubb, A.; Adolf, G.
 Eur. J. Haematol. 42, 270-275, 1989
 A:Title: Isolation and characterization of a tumor necrosis factor binding protein f
 A:Reference number: A60594; MUID:89171156; PMID:2924890
 A:Accession: A60594
 A:Molecule type: protein
 A:Residues: 41-43, 'X', 45-53, 'V', 55-57, 'XK', 60 <OLS>
 A:Experimental source: renal failure patient urine
 R:Engelmann, H.; Novick, D.; Wallach, D.
 J. Biol. Chem. 265, 1531-1536, 1990
 A:Title: Two tumor necrosis factor-binding proteins purified from human urine. Evid
 A:Reference number: A35010; MUID:90110215; PMID:2153136
 A:Accession: A35010
 A:Molecule type: protein
 A:Residues: 41-45 <ENG>
 A:Experimental source: normal urine
 R:Kajihara, J.; Asada, A.; Kirihara, S.; Kato, K.
 Biosci. Biotechnol. Biochem. 58, 2266-2268, 1994
 A:Title: Amino acid sequence of natural tumor necrosis factor alpha inhibitor purifi
 A:Reference number: J02404; MUID:95128033; PMID:7765720
 A:Accession: J02404
 A:Molecule type: protein
 A:Residues: 41-53, 'X', 55-144, 'X', 146-150, 'X', 152-186, 'X', 188-201 <KAJ>
 A:Experimental source: urine
 C:Comment: This protein is one of two known receptors for both TNF-alpha (cachectin)
 C:Genetics:
 A:Gene: GDB:TNFR1
 A:Cross-references: GDB:125913; OMIM:191190
 A:Map position: 12p13.2-12p13.2
 A:Introns: 13/3; 65/1; 108/1; 158/1; 184/2; 209/1; 247/1; 256/3; 353/1
 C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
 C:Keywords: duplication, glycoprotein, receptor, transmembrane protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-455/Product: tumor necrosis factor receptor 1 #status predicted <MAT>
 F:41-201/Domain: extracellular #status predicted <EXT>
 F:41-201/Product: TNF binding protein 1 (tumor necrosis factor alpha inhibitor) #sta
 F:44-82/Domain: NGF receptor repeat homology <NG1>
 F:84-126/Domain: NGF receptor repeat homology <NG2>
 F:127-167/Domain: NGF receptor repeat homology <NG3>

Gene 163, 263-266, 1995
 A:Title: Cloning of the cDNA encoding the porcine p55 tumor necrosis factor receptor.
 A:Reference number: J04302; MUID:96011645; PMID:7590278
 A:Accession: J04302
 A:Molecule type: mRNA
 A:Residues: 1-461 <SUT>
 A:Cross-references: GB:U19994; NID:g1141752; PIDN:AMC48499.1; PID:g1141753
 A:Accession: PC4093
 A:Molecule type: protein
 A:Residues: 1-7 <SUT>
 A:Experimental source: kidney cell line 15
 A:Genetics:
 C:Gene: tnfr
 C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
 C:Keywords: glycoprotein; kidney; receptor; transmembrane protein; tumor
 F:1-29/Domain: signal sequence #status predicted <Sig>
 F:30-461/Product: tumor necrosis factor receptor p55 #status predicted <MAT>
 F:44-194/Domain: extracellular cysteine rich #status predicted <EXT>
 F:84-82/Domain: NGF receptor repeat homology <NG1>
 F:84-126/Domain: NGF receptor repeat homology <NGF>
 F:211-231/Domain: transmembrane #status predicted <TM>
 F:361-447/Domain: signal transduction #status predicted <IT>
 F:54,145,151/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.9%; Score 244.5; DB 2; Length 461;
 Best Local Similarity 25.2%; Pred. No. 4.2e-07;
 Matches 105; Conservative 42; Mismatches 152; Indels 117; Gaps 22;

QY 7 TSAPLAMPPLV-----GLFGLAASQPOAVPVASENQCRCDEKEYEPOHR 56
 Db 4 STPELLP-PLVRLALVDVYPAGVHGLV-----LHPCGRKRESLCPQKYSHPQNR 55
 QY 57 -ICSCRCPPTGVYSAKC-SRIADYVCATCAENSYNENNYLTICQLCRPCDPPVMALEBIA 114
 Db 56 SICCKCKKRGYTLHNDCLGPGIDDCRCNDGNTFTASENHILFQCLSCSKSEMSQVETIS 115
 QY 115 PCTSKRKTQCRQCPGMFCANA---LECTHCELLSDCPGTE---AEIKDEVGKGNHNC 167
 Db 116 PCTYDRDVTGCGCRKNQYKRYMSETLFQCLNCSL---CPNGYVQLPCLKQDTI----- 165
 QY 168 VPCAGHGQNTSSPARQPHTRCENQGLVAAPESTAO-----SDTTCNPL----- 214
 Db 166 CNCHSGEFLR---DKECVSCVCKNADCKNLCPATSETRNDFODTGTVLPLVPIFFGL 221
 QY 215 -----EPLPEMSGSL--RRPOGEG-PNVVAGSEPPRAHYPFDVLP 258
 Db 222 CLAFFLFLVGLACRYQRMKPKIXSLICGSLTVKBEPEPLATA---PSFGP-----ITTF 273
 QY 259 LPISGDVSPVSTGLPAAPVLEAGVPQOQSPDLTREPQLEBEGSQVAHGTHGIVTGS 318
 Db 274 SPIP-SFSPPTTFSP-----VP-SFSPIS---SPTFTPCDMS----- 305
 QY 319 MTITGNITVYNGPVLAGPGPGDLPATPEP-----PYPIEEDDPGPGISTPHQ 368
 Db 306 -----NIKVISPPKEIAPPQAGPILPMPASTPVPPLPKWGSASHASAPAD 356

RESULT 9
 A42086
 CD30 antigen precursor - human
 N:Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30
 C:Species: Homo sapiens (man)
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000
 C:Accession: A42086
 R:Durrkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.
 Cell 68, 421-427, 1992
 A:Title: Molecular cloning and expression of a new member of the nerve growth factor rec
 A:Accession number: A42086; MUID:92154659; PMID:1310894
 A:Molecule type: mRNA
 A:Residues: 1-595 <DUR>
 A:Cross-references: GB:M83554; NID:g180095; PIDN:AAA51947.1; PID:g180096
 A:Experimental source: HUT-102 cell line

A>Note: sequence extracted from NCBI backbone (NCBIN:82088, NCBI:82090)
 C:Genetics:
 A:Gene: GDB:CD30; DIS166E
 A:Cross-references: GDB:131547; OMIM:153243
 A:Map position: 1p36-1p36
 C:Superfamily: NGF receptor repeat homology
 C:Keywords: glycoprotein; growth factor receptor; transmembrane protein
 F:1-18/Domain: signal sequence #status predicted <Sig>
 F:19-383/Domain: extracellular #status predicted <EXT>
 F:384-407/Domain: transmembrane #status predicted <TM>
 F:408-595/Domain: intracellular #status predicted <CYT>
 F:101,276/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.6%; Score 239; DB 2; Length 595;
 Best Local Similarity 25.1%; Pred. No. 1.1e-06;
 Matches 87; Conservative 29; Mismatches 129; Indels 102; Gaps 13;

QY 17 LVILGLFGLAASQPOAVPVASENQCRCDEKEYEPOHRICGSRCPPTGVYSAKSRIR 76
 Db 3 VILALGLIFGLALRAFQDRPFEDTCGNFSHTYDKAVKRCYRCRPGMLFPPTQCPQ-R 61
 QY 77 DTVG-ATCAENSYNENNYLTICQLCRPCDPPVMALEIAPCTSKRKTQCRQCPGMFCAAW 135
 Db 62 PTDCKRQCEPDYILDEADRCTACTVCSRD-----LVKTPCAWMSRVCECRPGMFCSTS 117
 QY 136 ALECHCELLSDCPPGTAEIKDEVGKGNHNCVPCAKGHFQNTSSPARQPHTRCENQ 195
 Db 118 AV-----NSC-----ARCFHSVCP-AG 134
 QY 196 LVEAPGAGSDTTCNPLEPLPEMSGSLKRRPOGEGP---NPVAGSEPPRAHYP 251
 Db 135 MIVAFPGTAOKNTVC---EPASFGVS-----PACASPNCKEPPSGITIPQAKPTP-- 181
 QY 252 PDLVQPLPLISGDVSPVSTGLPAAPVLEAGVPQOQSPDLTREPQLEBEGSQVAHGTHN- 310
 Db 182 -----VSPATSSASTMPVRGRTLAQEAASKILTRAPD-----SPSSVGRPSSD 224
 QY 311 -GHVGTGSMITGNITVYNGPVLAGPGPGDLPATPEPPIPIEEG 356
 Db 225 PGL-----SPTQPCPEGSDCKRQCEPDYILDEAG 254

RESULT 10
 S12783
 OX40 antigen precursor - rat
 N:Alternate names: nerve growth factor receptor homolog
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
 C:Accession: S12783; S08036
 R:Maliet, S.; Fossum, S.; Barclay, A.N.
 EMBO J. 9, 1063-1068, 1990
 A:Title: Characterization of the MRC OX40 antigen of activated CD4 positive T lympho
 A:Reference number: S12783; MUID:90214614; PMID:2157591
 A:Accession: S12783
 A:Molecule type: mRNA
 A:Residues: 1-271 <MAL>
 A:Cross-references: EMBL:X17037; NID:g57830; PIDN:CAA34897.1; PID:g57831
 C:Superfamily: CD27 antigen; NGF receptor repeat homology
 C:Keywords: growth factor receptor; transmembrane protein
 F:1-19/Domain: signal sequence #status predicted <Sig>
 F:20-271/Product: OX40 antigen #status predicted <MAT>
 F:211-235/Domain: transmembrane #status predicted <TM>

Query Match 10.3%; Score 232; DB 2; Length 271;
 Best Local Similarity 28.3%; Pred. No. 1.3e-06;
 Matches 66; Conservative 24; Mismatches 89; Indels 54; Gaps 8;

QY 48 KEYEPOHRICGSRCPPTGVYSAKSRIRDTVCATCAENSYNENNYLTICQLCRPCDPPV 107
 Db 28 KDITPSGCHK-CQRCQGHGAVSRCDHTRTVCHPCPEPGFYNEAVNYDT-CKQCTQCNRH 85
 QY 108 MGLEIAPCTSKRKTQCRQCPGMFCAMALECTHCELLSDCPTEAEIKDEVGKGNHNC 167

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:46:30 ; Search time 50.4702 Seconds

(without alignments)
760.006 Million cell updates/sec

Title: US-09-917-372-1

Perfect score: 2248

Sequence: 1 MLPMWATSAPGLAMGPLVLG.....EHCATPNSNGPRNQFTTHD 399

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2220	98.8	435	2	I54182	tumor necrosis fac
2	349.5	15.5	461	1	A35356	tumor necrosis fac
3	327.5	14.6	474	2	B38634	tumor necrosis fac
4	314.5	14.0	277	2	A60771	B-cell activation
5	312	13.9	459	2	I48854	gene murine tumour
6	267	11.9	651	2	JC7705	death receptor-6-
7	260	11.6	305	2	A46476	B cell-associated
8	244.5	10.9	461	2	JC4302	tumor necrosis fac
9	239	10.6	595	2	A42086	CD30 antigen precu
10	232	10.3	271	2	S12783	OX40 antigen precu
11	228	10.1	454	1	GOMST1	tumor necrosis fac
12	226	10.1	272	2	I48700	gene ox40 protein
13	222.5	9.9	455	1	GQHOT1	tumor necrosis fac
14	210	9.3	348	2	T28623	hypothetical prote
15	210	9.3	349	2	D36858	gene G4R protein -
16	209.5	9.3	277	2	I37552	OX40 homolog - hum
17	207	9.2	349	2	D72175	G2R protein - vari
18	204	9.1	461	1	GQRTT1	tumor necrosis fac
19	203.5	9.1	255	2	I38426	lymphocyte activat
20	196.5	8.7	425	1	A26431	nerve growth facto
21	191.5	8.5	493	2	JC5486	membrane glycoprot
22	189	8.4	325	2	B43692	T2 protein - rabbi
23	187.5	8.3	314	2	I37383	FAS soluble protei
24	185.5	8.3	256	2	B32393	T-cell antigen 4-1
25	183.5	8.2	326	1	GQVZML	T2 protein - myxom
26	183.5	8.2	335	2	A40036	apoptosis-mediatin
27	183.5	8.2	416	1	JN0006	nerve growth facto
28	175.5	7.8	427	1	GQHUN	nerve growth facto
29	162.5	7.2	327	2	A46484	apoptosis-mediatin

30	161.5	7.2	290	2	T24590	hypothetical prote
31	153	6.8	260	1	A46517	CD27 antigen precu
32	150	6.7	290	2	T24586	hypothetical prote
33	148.5	6.6	3635	2	T110053	laminin alpha 5 ch
34	148	6.6	303	2	T28999	hypothetical prote
35	146.5	6.5	1603	2	S23810	collagen alpha 1(X
36	145.5	6.5	324	2	JC2395	Fas antigen precu
37	145	6.5	1870	2	S37671	MHC class III hist
38	145	6.5	1872	2	S36152	MHC class III hist
39	144.5	6.4	1249	2	T14150	MHC class III hist
40	144	6.4	674	2	S13301	vesicle associated
41	143	6.4	2142	2	B35098	collagen alpha 1(X
42	142.5	6.3	1006	2	G86292	MHC class III hist
43	141.5	6.3	435	2	T15143	hypothetical prote
44	141.5	6.3	666	1	VCMVHL	hypothetical prote
45	140.5	6.2	2944	2	A54849	env polypeptid pr
						collagen alpha 1(V

ALIGNMENTS

RESULT 1

tumor necrosis factor receptor 2-related protein - human
C:Species: Homo sapiens (man)
C>Date: 24-May-1996 #sequence-revision 24-May-1996 #text-change 17-Mar-2000
C:Accession: I54182
R:Baens, M.; Chalfant, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.
Genomics 16, 214-218, 1993
A:Title: Construction and evaluation of a hncDNA library of human 12p transcribed se
A:Reference number: I54182; MUID:9325381; PMID:8486360
A:Accession: I54182
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-435 <RES>
A:Cross-references: GB:L04270; NID:g339761; PIDN:AAA36757.1; PID:g339762
C:Genetics:
A:Gene: GDB:LTRB
A:Cross-references: GDB:1230195; OMIM:600979
A:Map position: 12p13.3-12p13.1
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology

Query Match	98.8%	Score 2220	DB 2	Length 435
Best Local Similarity	91.7%	Pred. No. 8.4e-120		
Matches 399	0	Mismatches 0	Indels 36	Gaps 1
QY	1	MLPMWATSAPGLAMGPLVLGFLGLAASQPAVPPVASENOTCRDOKEYEPOHRI	60	
DB	1	MLPMWATSAPGLAMGPLVLGFLGLAASQPAVPPVASENOTCRDOKEYEPOHRI	60	
QY	61	RCPPGTIVYSAKCSRIKDYCATCAENSYEHMNYLTICQLCRPCDPVWGLTEIAPCTSKR	120	
DB	61	RCPPGTIVYSAKCSRIKDYCATCAENSYEHMNYLTICQLCRPCDPVWGLTEIAPCTSKR	120	
QY	121	KTCRCQCPGMCAMALECHTCELLSDCPPTBALKDEYKGNHVCPCAGHFOU	180	
DB	121	KTCRCQCPGMCAMALECHTCELLSDCPPTBALKDEYKGNHVCPCAGHFOU	180	
QY	181	PSACQPHTRCENGLVEAAGTASQSDTCNPLEPLPEKSGTMMALVLLPLAFILL	240	
DB	181	PSACQPHTRCENGLVEAAGTASQSDTCNPLEPLPEKSGTMMALVLLPLAFILL	240	
QY	223	-----GSLIKRRPQEGEPNPVAGSWEPPKAHPYEPDLVQPLLPISGD	264	
DB	241	ATVESCIMKSHPSLCRKIGSLIKRRPQEGEPNPVAGSWEPPKAHPYEPDLVQPLLPISGD	300	
QY	265	VSPVSTGLPAPVLEAGVPOQSPDLTRFQQLRPGSQVYAHGTNGIHYTGSMITTGN	324	
DB	301	VSPVSTGLPAPVLEAGVPOQSPDLTRFQQLRPGSQVYAHGTNGIHYTGSMITTGN	360	
QY	325	IYINGVVLGGPGLDPATPEPPYPIPEBGDGPGLSTPHQEDGKAMHLATETGCA	384	
DB	361	IYINGVVLGGPGLDPATPEPPYPIPEBGDGPGLSTPHQEDGKAMHLATETGCA	420	

Query Match	24.7%	Score 142	DB 2	Length 187998
Best Local Similarity	81.2%	Pred. No. 4,56-25		
Matches 177	Conservative 0	Mismatches 40	Indels 1	Gaps 1
QY 359	CCGGGGGCTCTAGTACTATCATCTGGCAACATCTATCATCTACANGAGCACTACTGGGGG	418		
Db 154396	CTCGACGCTCTGTACTGTCTACCGGGCATATCTACATATACAAATGGGCGAGTCTGGGGG	154455		
QY 419	GACCAACGGGCTCTGGAGACCTCCACAGTATCCCGGAACTCCATATCCCGGAAG	478		
Db 154456	GAAACACGGGGCTCTGGAGACCTCCAGTATCCCGCTGAGGCTTCATATCCCGAATCCCGAAG	154515		
QY 479	AGGGGAGACCTCTGGGCTCTCCGGGCTCTCTTACACCCACAGGAMAGATGGCAAGGCTTGGC	538		
Db 154516	AGGGAGAGGCTCTGGGCTCTCTGAGCTGTCTACACCTTACAGAGAGATGGCAAAAGCTTGGC	154575		
QY 539	ACCTAGCGGAGACAGAG-CAGCTGTGTGTCACACCTC	575		
Db 154576	ATCTGGTCTGAGACAGACACTAGGTGTCGAAGACCTC	154613		

RESULT 13

LOCUS	HUMTUMNEC	1605 bp	mRNA	linear	PRI 14-JAN-1999
DEFINITION	HUMTUMNEC				
	Homo sapiens (clone MCD18)	tumor necrosis factor	related		
	protein mRNA, complete exon and repeat region.				

ACCESSION	104489	104489.1	GI:340022	
VERSION	104489.1	GI:340022		
KEYWORDS	tumor necrosis factor receptor related protein.			
SOURCE	Human sapiens (tissue library: hncdna tc651) cDNA to mRNA.			
ORGANISM	Human sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Baens,M., Chaffanel,M., Cassiman,J.J., van den Berghe,H. and Marynen,P.			
TITLE	Construction and evaluation of a hncDNA library of human 12P transcribed sequences derived from a somatic cell hybrid			
JOURNAL	Genomics 16 (1), 214-218 (1993)			
MEDLINE	93252381			
PUBMED	8486360			
FEATURES	Location/Qualifiers			
SOURCE	1..1605			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/map="12p13"			
	/cell_line="M28 somatic cell hybrid"			
gene	/tissue_11b="hncdna tc651"			
	1..1605			
	/gene="TNFR"			
repeat_region	1..105			
	/partial			
	/note="putative"			
	/rpt_family="Alu"			
	/rpt_type="dispersed			
exon	250..444			
	/gene="TNFR"			
	/note="L04270 sequence homologue"			
BASE COUNT	472 a 394 c 323 g 416 t			
ORIGIN				
Query Match	14.4%; Score 83; DB 9; Length 1605;			
Best Local Similarity	100.0%; Pred. No. 3.1e-10;			
Matches	83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY	1 CTGTGTGAGGAGGAGCTCTCAGGACATGCCAGTCCGACACACACCTGCAAAATTCATTAGAG 60			
Db	363 CTGTGTGAGGAGGAGCTCTCAGGACATGCCAGTCCGACACACACCTGCAAAATTCATTAGAG 422			
OY	61 CCACGTGCCCCAGAGATGTCAAG 83			
Db	423 CCACGTGCCCCAGAGATGTCAAG 445			
RESULT 14				
HS27B9F/C				
LOCUS	HS27B9F 193 bp DNA linear PRI 19-OCT-1995			
DEFINITION	H.sapiens Cpg island DNA genomic MseI fragment, clone 27b9, forward			
ACCESSION	read cpq27b9.ftla.			
VERSION	Z60528			
KEYWORDS	Z60528.1 GI:1032632			
SOURCE	Cpg island; genomic MseI fragment.			
ORGANISM	Homo sapiens.			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Baens,M., Chaffanel,M., Cassiman,J.J., van den Berghe,H. and Marynen,P.			
TITLE	Construction and evaluation of a hncDNA library of human 12P transcribed sequences derived from a somatic cell hybrid			
JOURNAL	Genomics 16 (1), 214-218 (1993)			
MEDLINE	93252381			
PUBMED	8486360			
FEATURES	Location/Qualifiers			
SOURCE	1..1605			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/map="12p13"			
	/cell_line="M28 somatic cell hybrid"			
gene	/tissue_11b="hncdna tc651"			
	1..1605			
	/gene="TNFR"			
repeat_region	1..105			
	/partial			
	/note="putative"			
	/rpt_family="Alu"			
	/rpt_type="dispersed			
exon	250..444			
	/gene="TNFR"			
	/note="L04270 sequence homologue"			
BASE COUNT	472 a 394 c 323 g 416 t			
ORIGIN				
Query Match	14.4%; Score 83; DB 9; Length 1605;			
Best Local Similarity	100.0%; Pred. No. 3.1e-10;			
Matches	83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY	1 CTGTGTGAGGAGGAGCTCTCAGGACATGCCAGTCCGACACACCTGCAAAATTCATTAGAG 60			
Db	363 CTGTGTGAGGAGGAGCTCTCAGGACATGCCAGTCCGACACACCTGCAAAATTCATTAGAG 422			
OY	61 CCACGTGCCCCAGAGATGTCAAG 83			
Db	423 CCACGTGCCCCAGAGATGTCAAG 445			
RESULT 14				
HS27B9F				
LOCUS	HS27B9F 193 bp DNA linear PRI 19-OCT-1995			
DEFINITION	H.sapiens Cpg island DNA genomic MseI fragment, clone 27b9, forward			
ACCESSION	read cpq27b9.ftla.			
VERSION	Z60528			
KEYWORDS	Z60528.1 GI:1032632			
SOURCE	Cpg island; genomic MseI fragment.			
ORGANISM	Homo sapiens.			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Baens,M., Chaffanel,M., Cassiman,J.J., van den Berghe,H. and Marynen,P.			
TITLE	Construction and evaluation of a hncDNA library of human 12P transcribed sequences derived from a somatic cell hybrid			
JOURNAL	Genomics 16 (1), 214-218 (1993)			
MEDLINE	93252381			
PUBMED	8486360			
FEATURES	Location/Qualifiers			

Query Match	Best Local Similarity	28.0%;	Score 161;	DB 2;	Length 178228;
Matches 199;	Conservative	0;	Mismatches 45;	Indels 1;	Gaps 1;
332	CCCCAGTACCAATGAGGCAATGATCATGTCACCGGCGGGTCTATGATCATATCACTGGCAATCT	391	11	11	11
Db 152113	CCACAGGTGGCAATGGCAATGCACTGACCGGAGGCTGTGTGACGTGTACACCGCAATATCT	152054	11	11	11
QY	392	ACATCTACATATGACCACTACTGCGGGGACCAACCGGGTCTCTGGAGACTCCAGATACC	451	11	11
Db 152053	ACAAATATACAAATGGGCAATGCTGTGGGGGAGAACCGGGCTCTGGAGACCTCCAGTCCC	151994	11	11	11
QY	452	CCGAGACCTCCATTCACCCATTCCTCCGAAAGAGGGGAGACCTCTGGCCCTCCGGGCTCTTACAC	511	11	11
Db 151993	CTGAGCCCTCCATTCACCCGACTCCCGAAGAGGAGGAGCCCTCTGGCCCTCTGATGCTTACAC	151934	11	11	11
QY	512	CCCAACGAGAGATGGACGAGGCTTGGACCTAGCCGAGACAGAG-CACTGTGTGTCACA	570	11	11
Db 151933	CTTACACGAGAGATGGGCAAACTTGGCACTTGGCTGTGAGACAGAGACTAGGTGCCAAG	151874	11	11	11
QY	571	CCCTC 575	1111	1111	1111
Db 151873	ACCTC 151869	1111	1111	1111	1111
RESULT 12	AC125909	187998 bp	DNA	linear	HTG 23-JUL-2002
LOCUS	AC125909	Rattus norvegicus clone ch230-100N7, ***	SEQUENCING IN PROGRESS		
DEFINITION	AC125909	***, 64 unordered pieces.			
ACCESSION	AC125909				
VERSION	AC125909.1	GI:21671507			
KEYWORDS	HTG: HTGS_PHASE1.				
SOURCE	Rattus norvegicus				
ORGANISM	Rattus norvegicus				

```

repeat_region complement(17193..17563)
repeat_region /rpt_family="THEIA-int"
repeat_region complement(17564..17915)
repeat_region /rpt_family="THEIA"
repeat_region complement(18210..18342)
repeat_region /rpt_family="FLAM_A"
gene join(18391..18519,18744..18902,19603..19853)
gene /gene="Vesicle-associated membrane protein 1 (Syb1)"
repeat_region 19248..19275
repeat_region /rpt_family="(CA)n"
repeat_region 19276..19351
repeat_region /rpt_family="LIMBd"
repeat_region complement(19296..19475)
repeat_region /rpt_family="LIMCa"
repeat_region 19914..19933
repeat_region /rpt_family="(CCTG)n"
STS 20287..20651
STS /standard_name="GDB:214816"
STS 20290..20527
STS /standard_name="sts-M36200"
STS 20459..20650
STS /standard_name="sts-M36200"
STS 20841..21018
STS /standard_name="SHG-557"
STS 20971..21045
STS /standard_name="D12S111E"
STS 22216..22330
STS /standard_name="Cda1eb01"
STS 22518..22670
STS /standard_name="stSG4862"
gene complement(join(22595..22710,23806..23889,25797..26099,
27000..27338,31028..31297,31447..31677,32432..32495))

```

```

Query Match 41.6% Score 239 DB 9 Length 140026
Best Local Similarity 100.0% Pred. No. 2.7e+49
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 337 GGTACCATGGCATTCATGTCACCGGGGCTATGACTATACATGCAACATCTACATC 396
Db 94085 GGTACCATGGCATTCATGTCACCGGGGCTATGACTATACATGCAACATCTACATC 94026
QY 397 TACATGAGCAGTACTGAGGGGAGCACCAGGCTCTGAGAGCCTCCAGTACCCCGAA 456
Db 94025 TACATGAGCAGTACTGAGGGGAGCACCAGGCTCTGAGAGCCTCCAGTACCCCGAA 93966
QY 457 CCTCCATACCCCATTCCTCCGAGAGGAGACCTGCGCTCCCGGCTCTCTACACCCAC 516
Db 93965 CCTCCATACCCCATTCCTCCGAGAGGAGACCTGCGCTCCCGGCTCTCTACACCCAC 93906
QY 517 CAGGAGATGAGCAGGCTTGACCTGACGAGAGACGACGCTGTGTGACACCCCTC 575
Db 93905 CAGGAGATGAGCAGGCTTGACCTGACGAGAGACGACGCTGTGTGACACCCCTC 93847

```

```

RESULT 11
AC128082/c 178228 bp DNA linear HTG 19-JUL-2002
LOCUS Rattus norvegicus clone CH230-362C16, *** SEQUENCING IN PROGRESS
DEFINITION *** 64 unordered pieces.
ACCESSION AC128082
VERSION AC128082.1 GI:21908679
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 178228)

```

```

REFERENCE
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alshrooke,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

```

```

TITLE JOURNAL
REFERENCE DIRECT SUBMISSION
AUTHORS Unpublished
TITLE 2 (bases 1 to 178228)
JOURNAL Direct Submission
COMMENT Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZQW
Center clone name: CH230-362C16
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 116771 bases at least Q40
Consensus quality: 123909 bases at least Q30
Consensus quality: 129462 bases at least Q20

```

```

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 64 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1173: contig of 1173 bp in length
* 1174 1273: gap of unknown length
* 1274 2435: contig of 1162 bp in length
* 2436 2535: gap of unknown length

```


/translation="MRLPRASPCGLAMGBLLGLGLVASOPOLVPPYRIENQTC
DQREYEPHMDVCCSRKPGEFVAVCSRSODTVCCTCPHNSYNEHMHSLSTCOLCR
PCDIVLGEFEVAVCTSDRKAECRCQPMSCVYLDNEVCHEERLYLCOPGTAEVTD
EIMDTVNCVPRKPGHONRSPRARCIOIOLVIAAGTSTSDTICKNPPRP
GAMLLAIIILSYLFLFTTVILACAMMRHSICRKLGLTKRHPREBESPPCARPAD
PHFPLAEPLPMSGDLSPSPAGPTAPASLEEVYLDQOOSPLVQARELEAPGEGGYA
HGANGIHVTGGSVTVGNITLYNGPVLAGGTRGGRDPAPPEPPYPIPEBAGPSPSELS
TPYQEDGKAMHIAETETILGCDL"

BASE COUNT 316 a 537 c 455 g 306 t

ORIGIN

Query Match 48.5%; Score 278.8; DB 10; Length 1614;
Best Local Similarity 80.0%; Pred. No. 3.9e-59;
Matches 328; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 146 CTCGAAGGCCCATCATCTCCCTGACTGATACAGCCCATCTACCCATTTCTGAG 205
Db 1000 CTCGAAGGCCCATCATCTCCCTGACTGATACAGCCCATCTACCCATTTCTGAG 1059

QY 206 ATGTTTCCCAATATCCACTGAGGCTCCCGCAGCCCATTTTGGAGGAGGCTGCCG 265
Db 1060 ACTGTCCCATCCCTCTGCTGACCCCAAGGCGCCCTCTTGGAGAAATGGGTCTAC 1119

QY 266 AACAGCAGAGTCTCTGACCTGACAGAGGAGCCGCAATTTGGAACCCGGGAGACAGACC 325
Db 1120 AACAGCAGAGTCTCTGCTGACAGGAGGAGGCTGGAACCTTGGGAGAACATGGCC 1179

QY 326 AGGTGGCCCAAGTACCATGATGATGACGGGGGCTATGACTATCACTGACGA 385
Db 1180 AGGTGGCCCAAGTACCATGATGATGACGGGGGCTCTGTGACTGTACCGGCA 1239

QY 386 ACATCTACATCTACATATGAGACAGTACTGAGGGGAGACACCGGGTCTTGAGACCTCCAG 445
Db 1240 ACATCTACATCTACATATGAGGAGGAGTCTGAGGGGAGACACCGGGTCTTGAGACCTCCAG 1299

QY 446 CTACCCCGGAACCTTCATATCCCATTCGGAAGAGGGGAGCCCTGCGCCCTCCCGGGCTCT 505
Db 1300 CTCCCGCCGAGGCTTCATATCCCATTCGGAAGAGGGGAGGCTCTGCGCCCTCTGAGCTGT 1359

QY 506 CTACACCCCGAGGAGATGAGCAAGGCTTGACACCTGAGGAGGAGACAGAG 555
Db 1360 CTACACCCCGAGGAGATGAGCAAGGCTTGACATCTGCGGAGACAGAG 1409

RESULT 9
MMU29173 2076 bp mRNA linear ROD 28-JUN-1995
DEFINITION Mus musculus lymphotoxin-beta receptor mRNA, complete cds.
ACCESSION U29173
VERSION U29173.1 GI:881620
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2076)
AUTHORS Force, W.R., Williams-Abbott, L., Browning, J., Hession, C., Tizard, R.
and Ware, C.F.
TITLE Cloning and Characterization of the Mouse Lymphotoxin-beta Receptor
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2076)
AUTHORS Force, W.R.
TITLE Direct Submission
JOURNAL Submitted (15-JUN-1995) Walker R. Force, Biomedical Sciences, U.C.
Riverside, Riverside, CA 92521-0121, USA
FEATURES
Location/Qualifiers
1..2076
/organism="Mus musculus"
/strain="CWB"
/db_xref="taxon:10090"
/sex="female"
/tissue_type="lung"
170..1417
CDS

/note="transmembrane protein"
/codon_start=1
/product="Lymphotoxin-beta receptor"
/protein_id="AA68964.1"
/db_xref="GI:881621"

/translation="MRLPRASPCGLAMGBLLGLGLVASOPOLVPPYRIENQTC
DQREYEPHMDVCCSRKPGEFVAVCSRSODTVCCTCPHNSYNEHMHSLSTCOLCR
PCDIVLGEFEVAVCTSDRKAECRCQPMSCVYLDNEVCHEERLYLCOPGTAEVTD
EIMDTVNCVPRKPGHONRSPRARCIOIOLVIAAGTSTSDTICKNPPRP
GAMLLAIIILSYLFLFTTVILACAMMRHSICRKLGLTKRHPREBESPPCARPAD
PHFPLAEPLPMSGDLSPSPAGPTAPASLEEVYLDQOOSPLVQARELEAPGEGGYA
HGANGIHVTGGSVTVGNITLYNGPVLAGGTRGGRDPAPPEPPYPIPEBAGPSPSELS
TPYQEDGKAMHIAETETILGCDL"

BASE COUNT 436 a 652 c 597 g 391 t

ORIGIN

Query Match 48.5%; Score 278.8; DB 10; Length 2076;
Best Local Similarity 80.0%; Pred. No. 3.9e-59;
Matches 328; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 146 CTCGAAGGCCCATCATCTCCCTGACTGATACAGCCCATCTACCCATTTCTGAG 205
Db 984 CTCGAAGGCCCATCATCTCCCTGACTGATACAGCCCATCTACCCATTTCTGAG 1043

QY 206 ATGTTTCCCAATATCCACTGAGGCTCCCGCAGCCCATTTTGGAGGAGGCTGCCG 265
Db 1044 ACTGTCCCATCCCTCTGCTGACCCCAAGGCGCCCTCTTGGAGAAATGGGTCTAC 1103

QY 266 AACAGCAGAGTCTCTGACCTGACAGAGGAGCCGCAATTTGGAACCCGGGAGACAGACC 325
Db 1104 AACAGCAGAGTCTCTGCTGACAGGAGGAGGCTGGAACCTTGGGAGAACATGGCC 1163

QY 326 AGGTGGCCCAAGTACCATGATGATGACGGGGGCTATGACTATCACTGACGA 385
Db 1164 AGGTGGCCCAAGTACCATGATGATGACGGGGGCTCTGTGACTGTACCGGCA 1223

QY 386 ACATCTACATCTACATATGAGACAGTACTGAGGGGAGACACCGGGTCTTGAGACCTCCAG 445
Db 1224 ACATCTACATCTACATATGAGGAGGAGTCTGAGGGGAGACACCGGGTCTTGAGACCTCCAG 1283

QY 446 CTACCCCGGAACCTTCATATCCCATTCGGAAGAGGGGAGCCCTGCGCCCTCCCGGGCTCT 505
Db 1284 CTCCCGCCGAGGCTTCATATCCCATTCGGAAGAGGGGAGGCTCTGCGCCCTCTGAGCTGT 1343

QY 506 CTACACCCCGAGGAGATGAGCAAGGCTTGACACCTGAGGAGGAGACAGAG 555
Db 1344 CTACACCCCGAGGAGATGAGCAAGGCTTGACATCTGCGGAGACAGAG 1393

RESULT 10
AC005840 140026 bp DNA linear PRI 20-OCT-2000
DEFINITION Homo sapiens complete sequence of a PAC clone RPI-102E24 containing
SYB1, CD27, and SCNN1A genes.
ACCESSION AC005840
VERSION AC005840.2 GI:10938025
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 140026)
AUTHORS Montgomery, K.T., Lau, S.T. and Kucherlapati, R.
TITLE High throughput Sequencing of Human Chromosome 12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 140026)
AUTHORS Montgomery, K.T., Lau, S.T. and Kucherlapati, R.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1998) Department of Molecular Biology, Albert
Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY
10461, USA
REFERENCE 3 (bases 1 to 140026)
AUTHORS Montgomery, K.T., Lau, S.T. and Kucherlapati, R.

PUBMED	8486360	Location/Qualifiers
FEATURES		1. 2136
SOURCE		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/tissue_type="liver"
		/tissue_lib="liver CDNA of P.M."
		169..1476
		/note="putative"
		/codon_start=1
		/product="tumor necrosis factor receptor 2 related protein"
		/protein_id="AA36757.1"
		/db_xref="GI:339762"
		/translation="MLPFAATSAAGLAWGLVGLFEILAAASQPAVPPASAEQTR DQERYIEPQRIICCSRCPPGITYVASACTIRIVCTACVENSEMWNVLITTCOLR PCDVGLEIEIAPCTSRKRYQCRQPMFCFAAVALICTHEELLSDCPGETAELEKDEV KSGNNHCPCPKAGHFQMTSSPSASACOPHTCEENGLVEAAQASDTCKNPLELP PEKSGTMIAMLAVALPLFLFLAVFSGCIKMSHSLCRKASGLIKRRDQEGPYVAG SWEPKRAYPEFDPLOPLFLISGDSVSVTSGLPAAVYLEAGVPOQSPDLTRPOLE PESGKAYAHGTNGIHVYGLSMETITGNLYITNGPLVGGPQPGDLPARPEPPYPIPEG DPRPGLSTPHQEDGKAMHLAEFHCATFTSNKRPQFTTHD"
BASE COUNT	446 a 706 c 608 g 376 t	
ORIGIN		
Query Match	85.9%; Score 494; DB 9; Length 2136;	
Best Local Similarity	100.0%; Pred. No. 6.9e-113;	
Matches	494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	82	GGATGCTGCTCAGAGCGCTCCGAGGAGAGGAGACCCAACTCTGTAGCTGGAAAGCTGG 141
Db	943	GGATCGCTGCTCAGAGCGCTCCGAGGAGAGGAGACCCAACTCTGTAGCTGGAAAGCTGG 1002
QY	142	GAGCCTCCGAGAGGCCATCATCTACTCCCTGACTGTGATACGCCATCTGCTACCATTTCT 201
Db	1003	GAGCCTCCGAGAGGCCATCATCTACTCCCTGACTGTGATACGCCATCTGCTACCATTTCT 1062
QY	202	GGAGATGTTTCCCGCAGTATACCTGGGCTCCCGCAGGCCCCAGATTTTGGAGGCAAGGGGTG 261
Db	1063	GGAGATGTTTCCCGCAGTATACCTGGGCTCCCGCAGGCCCCAGATTTTGGAGGCAAGGGGTG 1122
QY	262	CCGCAACGACGAGAGCTCTGGACCTGACACGAGGAGCCGCACTTGGAAACCCGGGGAGCAG 321
Db	1123	CCGCAACGACGAGAGCTCTGGACCTGACACGAGGAGCCGCACTTGGAAACCCGGGGAGCAG 1182
QY	322	AGCCAGTGGGCCAAGGTATCAATGGCAATGGCAATTAATGTCACCGCGGGCTGATGACTATCACT 381
Db	1183	AGCCAGTGGGCCAAGGTATCAATGGCAATGGCAATTAATGTCACCGCGGGCTGATGACTATCACT 1242
QY	382	GGCAACATCTACATCTACAAATGGACAGTACTGGGGGACCAACCGGGTCTTGGAGACCTC 441
Db	1243	GGCAACATCTACATCTACAAATGGACAGTACTGGGGGACCAACCGGGTCTTGGAGACCTC 1302
QY	442	CCAGCTACCCCGGAACCTCCATACCCCAATTCGCCGAAGAAGGGGAGACCTGGGCCCTCCGGG 501
Db	1303	CCAGCTACCCCGGAACCTCCATACCCCAATTCGCCGAAGAAGGGGAGACCTGGGCCCTCCGGG 1362
QY	502	CTCTGTACACCCACACAGAGAGATGGCAAGGCTTGGCACTAGCGGAGACAGACACTGT 561
Db	1363	CTCTGTACACCCACACAGAGAGATGGCAAGGCTTGGCACTAGCGGAGACAGACACTGT 1422
QY	562	GGTGCACACCCCTC 575
Db	1423	GGTGCACACCCCTC 1436
RESULT 6		
LOCUS	AK027080	
DEFINITION	AK027080 2148 bp mRNA linear PRI 29-SEP-2000	
	Homo sapiens cDNA: FLJ234427 fls, clone HRC04768, highly similar to	
	HUMTNFRP Homo sapiens tumor necrosis factor receptor 2 related	
	protein mRNA.	
ACCESSION	AK027080	

VERSION	AK027080.1	GI:10440111
KEYWORDS	Oligo capping; fls (full insert sequence).	
SEQUENCE	Homo sapiens primary human renal epithelial cells cDNA to mRNA, clone_11b:HRC clone:HRC04788.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (siles) Kawabata,A., Hiki,I., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shihabara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.	
TITLE	NEO human cDNA sequencing project	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 2148)	
AUTHORS	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shihabara,T., Tanaka,T. and Nakamura,Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (23-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:shirokane.utsu.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)	
COMMENT	NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- 6' 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).	
FEATURES	Location/Qualifiers	
source	1..2148 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="HRC04788" /cell_type="primary human renal epithelial cells" /clone_11b="HRC" /note="Cloning vector pME18SF13"	
misc-feature	1..2148 /note="highly similar to HUNTNRFP Homo sapiens tumor necrosis factor receptor 2 related protein mRNA"	
BASE COUNT	462 a 705 c 606 g 375 t	
ORIGIN		
Query Match	85.9%; Score 494; DB 9; Length 2148;	
Best Local Similarity	100.0%; Pred. No. 6.9e-113;	
Matches 494; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	82	GGATGGCTGCTCAAGAGGGCTCGGAGGAGGAGGAGGCCAATCCGTGAGTGAAGCTGG 141
DB	966	GGATGGCTGCTCAAGAGGGCTCGGAGGAGGAGGAGGCCAATCCGTGAGTGAAGCTGG 1025
QY	142	GAGCTCCGAAAGGCCCATCATCTCCCTCGACTGTGTACAGCCACTGTATCCCATTTCT 201
DB	1026	GAGCTCCGAAAGGCCCATCATCTCCCTCGACTGTGTACAGCCACTGTATCCCATTTCT 1085
QY	202	GGAGATGTTTCCCAAGTATCCACTGGGCTCCCGCAGGCCCGATTGGAGGACAGGGTG 261
DB	1086	GGAGATGTTTCCCAAGTATCCACTGGGCTCCCGCAGGCCCGATTGGAGGACAGGGTG 1145
QY	262	CCGCAACAGCAGAGTCCTGTGGAGCTGACACGAGAGCCGACATTTGGAACCCGGAGGAC 321
DB	1146	CCGCAACAGCAGAGTCCTGTGGAGCTGACACGAGAGCCGACATTTGGAACCCGGAGGAC 1205
QY	322	AGCCAGGTGGCCACGATACCAATGGCATTTGTCACCGGGGGCTATATGACTATCACT 381
DB	1206	AGCCAGGTGGCCACGATACCAATGGCATTTGTCACCGGGGGCTATATGACTATCACT 1265
QY	382	GGCAACATTCATATCATATGACATGACACAGTACAGGGGGGACACCGGGGTCCTGGAACCTC 441
DB	1266	GGCAACATTCATATCATATGACATGACACAGTACAGGGGGGACACCGGGGTCCTGGAACCTC 1325
QY	442	CCAGGTACCCCGGAAGCTTCATACCCCATTTCCCGAAGAGGGGAGCTGAGCCCTCCGGG 501

REFERENCE	TITLE	JOURNAL	FEATURES	BASE COUNT	ORIGIN
AX332212	AX332212.1	GI:18122846			
KEYWORDS	human.				
ORGANISM	Homo sapiens				
AUTHORS	Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrijan, S., Soppel, D.R. and Weaver, Z.				
TITLE	Cancer gene determination and therapeutic screening using signature gene sets				
JOURNAL	Patent: WO 0194629-A 2721 13-DEC-2001;				
FEATURES	Avalon Pharmaceuticals (US)				
SOURCE	Location/Qualifiers				
BASE COUNT	446 a	706 c	608 g	376 t	
ORIGIN	db_xref="taxon:9606"				
Query Match	85.9%; Score 494;	DB 6;	Length 2136;		
Best Local Similarity	100.0%;	Pred. No. 6.9e-113;			
Matches 494;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	82	GGATGGCTGTCAGAGGCGTCGCCGAGGAGAGAGACCACATTCCTGTAGCTGGAGAGCTGG	141		
Db	943	GGATGGCTGTCACAGAGGCGTCGCCGAGGAGAGAGACCACATTCCTGTAGCTGGAGAGCTGG	1002		
QY	142	GAGCCTCCGAGAGGCGCCATCCATCACTTCCCTGACTGTGTACAGCCACTGCTACCCATTTCT	201		
Db	1003	GAGCCTCCGAGAGGCGCCATCCATCACTTCCCTGACTGTGTACAGCCACTGCTACCCATTTCT	1062		
QY	202	GGAGATGTTTCCCGAGTATTCACCTGGGCTCCCGCAGCCCGAGTTTGGAGGACAGGGGTG	261		
Db	1063	GGAGATGTTTCCCGAGTATTCACCTGGGCTCCCGCAGCCCGAGTTTGGAGGACAGGGGTG	1122		
QY	262	CCGCAACGACGAGAGTCCCTGTGGACTTACACAGAGGAGCCGCACTTGGAAACCCGGGAGCAG	321		
Db	1123	CCGCAACGACGAGAGTCCCTGTGGACTTACACAGAGGAGCCGCACTTGGAAACCCGGGAGCAG	1182		
QY	322	AGCCAGGTGGGCCACGAGTACCAATGGCAATTCATGTGCACCGGCGGGTCTATGACTTCACT	381		
Db	1183	AGCCAGGTGGGCCACGAGTACCAATGGCAATTCATGTGCACCGGCGGGTCTATGACTTCACT	1242		
QY	382	GGCAACATCTACATCTACCAATGGACCACTACTGGGGGAGACACCGGGTCTGGAGACCTC	441		
Db	1243	GGCAACATCTACATCTACCAATGGACCACTACTGGGGGAGACACCGGGTCTGGAGACCTC	1302		
QY	442	CCAGGTACCCCGGAACCTTCATACCCCATTTCCCGAAGAGGGGAGACCTTGGCCCTCCGGG	501		
Db	1303	CCAGGTACCCCGGAACCTTCATACCCCATTTCCCGAAGAGGGGAGACCTTGGCCCTCCGGG	1362		
QY	502	CTCTCTACACCCCGACAGAGATGGCAAGGGCTTGGCACTTAGGGAGAGACAGACACTGT	561		
Db	1363	CTCTCTACACCCCGACAGAGATGGCAAGGGCTTGGCACTTAGGGAGAGACAGACACTGT	1422		
QY	562	GGTGGCACACCCCTC	575		
Db	1423	GGTGGCACACCCCTC	1436		
RESULT 4	AX409488	2136 bp	DNA	linear	PAT 14-JUN-2002
LOCUS	AX409488				
DEFINITION	Sequence 2135 from Patent WO0229103.				
ACCESSION	AX409488				
VERSION	AX409488.1	GI:21442193			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				

REFERENCE						
AUTHORS	TITLE	JOURNAL	GENE LOGIC INC (US)	Location/Qualifiers	1..2136	
FEATURES						
SOURCE	/organism="Homo sapiens"	/db_xref="taxon:9606"	/note="EMBL/GenBank Accession No.	L04270"		
BASE COUNT	446 a	706 c	608 g	376 t		
ORIGIN						
Query Match 85.9%; Score 494; DB 6; Length 2136;						
Best Local Similarity 100.0%; Prid. No. 6.9e-113;						
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
OY	82	GGAATCGCTGCTCAAGAGGCGTCGCCGAGGAGAGGGACCACAACTCCTGTAGTGCAGAAGCTGG	141			
Dd	943	GGATTCGCTGCTCAGAGGCGCTCCGAGGAGAGGAGAACCCAATTCTTGTAAGTAGACTGG	1002			
OY	142	GAGCCTCCGAAGGCCCATCATCTACTTCCCCTGACTTTGGTAGACCACTGCTAOCCTACT	201			
Dd	1003	GAGCCTCCGAAGGCCCATCATCTACTTCCCCTGACTTTGGTAGACCACTGCTAOCCTACT	1062			
OY	202	GGAGATGTTTTCCCCAGTATCCACTGSGCTCCCGCAGGCCCAAGTTTGGAGGCAAGGGTG	261			
Dd	1063	GGAGATGTTTTCCCCAGTATCCACTGSGCTCCCGCAGGCCCAAGTTTGGAGGCAAGGGTG	1122			
OY	262	CCGCAACAGCAGAGTCCCTTGACCTGACACAGGAGAGCGCGAGTTGGAAACCCGGGAGGAG	321			
Dd	1123	CCGCAACAGCAGAGTCCCTTGACCTGACACAGGAGAGCGCGAGTTGGAAACCCGGGAGGAG	1182			
OY	322	AGCCAGGTGGGCCAAGGTATCCATGGAATGGAATGATGTCACCGGGCTCATGACTATCACT	381			
Dd	1183	AGCCAGGTGGGCCAAGGTATCCATGGAATGGAATGATGTCACCGGGCTCATGACTATCACT	1242			
OY	382	GGCAACATCTCATCTTACATATGAGACAGTACTGGGGGACACAAGGGTCTTGAGAGCTC	441			
Dd	1243	GGCAACATCTCATCTTACATATGAGACAGTACTGGGGGACACAAGGGTCTTGAGAGCTC	1302			
OY	442	CCAGTATACCCCCGAAACCTCCATACCCCATTCCCGGAAAGGGGGAGACCTGGGCTCCCGGG	501			
Dd	1303	CCAGTATACCCCCGAAACCTCCATACCCCATTCCCGGAAAGGGGGAGACCTGGGCTCCCGGG	1362			
OY	502	CCTCTCTACACCCCAACAGAGATGAGGCAAGGCTTGAGCACCTAGCGGAGACAGAGCACTGT	561			
Dd	1363	CCTCTCTACACCCCAACAGAGATGAGGCAAGGCTTGAGCACCTAGCGGAGACAGAGCACTGT	1422			
OY	562	GGTGGCACACCCCTC	575			
Dd	1423	GGTGGCACACCCCTC	1436			
RESULT 5						
HUMTNFRFP	LOCUS	HUMTNFRFP				
DEFINITION	HOMO SAPIENS (clone CD18) tumor necrosis factor receptor 2 related protein mRNA, complete cds.					
ACCESSION	L04270					
VERSION	L04270.1 GI:339761					
KEYWORDS	tumor necrosis factor receptor 2 related protein.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (bases 1 to 2136)					
AUTHORS	Buens,M., Chaffanel,M., Cassiman,J.J., van den Berghe,H. and Marynen,P.					
TITLE	Construction and evaluation of a hncDNA library of human 12p transcribed sequences derived from a somatic cell hybrid					
JOURNAL	Genomics 16 (1), 214-218 (1993)					
MEDLINE	93252381					

Db	1410	GGTGGCACACCCCTC	1423	
RESULT 2				
LOCUS	AX331906		2136 bp	DNA
DEFINITION	Sequence 2415 from Patent WO0194629.			linear
ACCESSION	AX331906			
VERSION	AX331906.1			
KEYWORDS	GI:18122540			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS	1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horigan, S., Soppel, D.R. and Weaver, Z.			
TITLE	Cancer gene determination and therapeutic screening using signature gene sets			
JOURNAL	Patent: WO 0194629-A 2415 13-DEC-2001;			
FEATURES	Avalon Pharmaceuticals (US)			
source	Location/Qualifiers			
	1. 2136			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
BASE COUNT	446 a 706 c 608 g 376 t			
ORIGIN				
Query Match	85.98; Score 494; DB 6; Length 2136;			
Best Local Similarity	100.0%; Pred. No. 6.9e-113;			
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY	82	GGATGCTGCTCAAGAGGCGTCCGACGAGGAGAGGAGCCCAATCTGTAGCTGAAGCTGG	141	
Db	943	GGATGCTGCTCAAGAGGCGTCCGACGAGGAGAGGAGCCCAATCTGTAGCTGAAGCTGG	1002	
OY	142	GAGCCTCCGAAGGCCCATTCATCTTCCCTGACTTGGTACAGCCACATGCTACCATTTCT	201	
Db	1003	GAGCCTCCGAAGGCCCATTCATCTTCCCTGACTTGGTACAGCCACATGCTACCATTTCT	1062	
OY	202	GGAGATGTTTCCGCCAGTATTCACCTGGGCTCCCGACGCCAGTTTGGAGGACAGGGTG	261	
Db	1063	GGAGATGTTTCCGCCAGTATTCACCTGGGCTCCCGACGCCAGTTTGGAGGACAGGGTG	1122	
OY	262	CCGCAACAGCAGAGTCTCTGTGACCTGACACAGGAGCCGCACTTGGAAACCCGGGAGCAG	321	
Db	1123	CCGCAACAGCAGAGTCTCTGTGACCTGACACAGGAGCCGCACTTGGAAACCCGGGAGCAG	1182	
OY	322	AGCCAGGTGGCCCAAGGTATCCATATGGCATTTATGTACACCGGGGGTCTATGACTATCACT	381	
Db	1183	AGCCAGGTGGCCCAAGGTATCCATATGGCATTTATGTACACCGGGGGTCTATGACTATCACT	1242	
OY	382	GGCAACATCTACATTAATGAGACCACTAGCTGGGGGAGACACCGGCTCTGGAAGACTC	441	
Db	1243	GGCAACATCTACATTAATGAGACCACTAGCTGGGGGAGACACCGGCTCTGGAAGACTC	1302	
OY	442	CCAGTACCCCGGAAGCTTCATACCCCAATTCGCCGAAGAGGGGAGCCCTGCTCCGGG	501	
Db	1303	CCAGTACCCCGGAAGCTTCATACCCCAATTCGCCGAAGAGGGGAGCCCTGCTCCGGG	1362	
OY	502	CTCTTACACCCACAGGAAGATGGCAAGCTTGGCACTTGGCGGAGACAGACACTGT	561	
Db	1363	CTCTTACACCCACAGGAAGATGGCAAGCTTGGCACTTGGCGGAGACAGACACTGT	1422	
OY	562	GGTGGCACACCCCTC	575	
Db	1423	GGTGGCACACCCCTC	1436	
RESULT 3				
LOCUS	AX332212		2136 bp	DNA
DEFINITION	Sequence 2721 from Patent WO0194629.			linear
PAT	09-JAN-2002			

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:04:15 ; Search time 1072.25 seconds
(without alignments)
15606.600 Million cell updates/sec

Title: US-09-917-372-2_COPY_804_1378

Perfect score: 575
Sequence: 1 ctgtgtgagcagctccagg.....cactgtgtgtccacaccctc 575

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:**
1: gb_da:**
2: gb_htg:**
3: gb_in:**
4: gb_cm:**
5: gb_ov:**
6: gb_pat:**
7: gb_ph:**
8: gb_pl:**
9: gb_pr:**
10: gb_ro:**
11: gb_sts:**
12: gb_sy:**
13: gb_un:**
14: gb_vl:**
15: gb_da:**
16: em_fun:**
17: em_hum:**
18: em_in:**
19: em_mu:**
20: em_om:**
21: em_or:**
22: em_ov:**
23: em_pat:**
24: em_ph:**
25: em_pl:**
26: em_ro:**
27: em_sts:**
28: em_un:**
29: em_vl:**
30: em_htg_hum:**
31: em_htg_inv:**
32: em_htg_other:**
33: em_htg_mus:**
34: em_htg_pin:**
35: em_htg_rnd:**
36: em_htg_mam:**
37: em_htg_vrt:**
38: em_sy:**
39: em_htgo_hum:**
40: em_htgo_mus:**
41: em_htgo_other:**

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	494	85.9	2091	9	AK095418	AK095418 Homo sapi
2	494	85.9	2136	6	AX331906	AX331906 Sequence
3	494	85.9	2136	6	AX332212	AX332212 Sequence
4	494	85.9	2136	6	AX409488	AX409488 Sequence
5	494	85.9	2136	6	AX409488	AX409488 Sequence
6	494	85.9	2136	6	AX409488	AX409488 Sequence
7	494	85.9	2136	6	AX409488	AX409488 Sequence
8	494	85.9	2136	6	AX409488	AX409488 Sequence
9	494	85.9	2136	6	AX409488	AX409488 Sequence
10	494	85.9	2136	6	AX409488	AX409488 Sequence
11	494	85.9	2136	6	AX409488	AX409488 Sequence
12	494	85.9	2136	6	AX409488	AX409488 Sequence
13	494	85.9	2136	6	AX409488	AX409488 Sequence
14	494	85.9	2136	6	AX409488	AX409488 Sequence
15	494	85.9	2136	6	AX409488	AX409488 Sequence
16	494	85.9	2136	6	AX409488	AX409488 Sequence
17	494	85.9	2136	6	AX409488	AX409488 Sequence
18	494	85.9	2136	6	AX409488	AX409488 Sequence
19	494	85.9	2136	6	AX409488	AX409488 Sequence
20	494	85.9	2136	6	AX409488	AX409488 Sequence
21	494	85.9	2136	6	AX409488	AX409488 Sequence
22	494	85.9	2136	6	AX409488	AX409488 Sequence
23	494	85.9	2136	6	AX409488	AX409488 Sequence
24	494	85.9	2136	6	AX409488	AX409488 Sequence
25	494	85.9	2136	6	AX409488	AX409488 Sequence
26	494	85.9	2136	6	AX409488	AX409488 Sequence
27	494	85.9	2136	6	AX409488	AX409488 Sequence
28	494	85.9	2136	6	AX409488	AX409488 Sequence
29	494	85.9	2136	6	AX409488	AX409488 Sequence
30	494	85.9	2136	6	AX409488	AX409488 Sequence
31	494	85.9	2136	6	AX409488	AX409488 Sequence
32	494	85.9	2136	6	AX409488	AX409488 Sequence
33	494	85.9	2136	6	AX409488	AX409488 Sequence
34	494	85.9	2136	6	AX409488	AX409488 Sequence
35	494	85.9	2136	6	AX409488	AX409488 Sequence
36	494	85.9	2136	6	AX409488	AX409488 Sequence
37	494	85.9	2136	6	AX409488	AX409488 Sequence
38	494	85.9	2136	6	AX409488	AX409488 Sequence
39	494	85.9	2136	6	AX409488	AX409488 Sequence
40	494	85.9	2136	6	AX409488	AX409488 Sequence
41	494	85.9	2136	6	AX409488	AX409488 Sequence

ALIGNMENTS

RESULT 1
AK095418
LOCUS AK095418
DEFINITION Homo sapiens CDNA FLJ138099 fls, clone D3OST1000238, highly similar
to LYMPHOXIN-BETA RECEPTOR PRECURSOR.
ACCESSION AK095418
VERSION AK095418.1 GI:21754669
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens CD34+ Cells cDNA to mRNA, clone D3OST1
clone.D3OST1000238.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Oshima,A., Takahashi-Fuji,I., Tanase,T., Imose,N., Takeuchi,K.,
AUTHORS


```
QY 197 ACTGCCAGTCCGACACACCTGCAAAAATCCATTAGAGCCACTGCCCCAGATGTCA 256
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 452 ACCTCTACTCTTGACACATCTGTAAATCCACCCGAGGACAGGAAATGCTGCTA 511
QY 257 GGATCGCTGCT 267
      || | | | |
Db 512 CTAGCCATCCT 522

RESULT 14
US-09-907-372-13
; Sequence 13, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 206
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 700302531H1
US-09-907-372-13

Query Match          20.9%; Score 135.8; DB 10; Length 206;
Best Local Similarity 81.0%; Pred. No. 1.1e-32;
Matches 158; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 455 CTGGACCTGACACGAGGCCGAGTTGGAACCCGGGGAGCAGAGCGGTGCCCCAGGT 514
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2 CTGATCCAGGCCAGGGAGCTGAGAGCTGAGCCTGGGGAAATGCGCGCCAGGT 61
QY 515 ACCAATGGCATTTATGTACCGCGGGGTCTATGACTATCACTGCAACATCTACATCTAC 574
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 GCGAATGGCATTTACCTACCGGAGGCTCTGTACTGTACCGCAATATCTACATATAC 121
QY 575 AATGACACGATGCTGGGGGAGCCAGCGGTCGAGAGACTCCAGCTACCCGGAACCT 634
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 AATGGCCAGTGTCTGGGGGAGACAGGGGCCCTGAGAGACCTCCAGCTCCCTGAGCT 181
QY 635 CCATACCCCATTCCTCC 649
      || | | | | | | | | | |
Db 182 CCATACCCGACCTCC 196

RESULT 15
US-09-907-372-15
; Sequence 15, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 702022948H1
US-09-907-372-15
```

```
Query Match          20.9%; Score 135.8; DB 10; Length 471;
Best Local Similarity 76.3%; Pred. No. 1.3e-32;
Matches 167; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 17 GACCTACTTTTGTGACTGCCCCGCTGGCAGTGAAGCCGAGCTCAAGATGAAGTTGGGAAG 76
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 205 GAGCGGCTTGTACTCTGACGCGCCTGGCAGAAAGCTGAGGTCAAGATGAATATGAGAT 264
QY 77 GGTACACACACTGCGTCCCTCCGAAAGGACGAGGCACTTCCAGAAATCCTCTCCCGAGC 136
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 265 ACTGAAGTCAACTGTGTCCCTCTAAGCCAGACACTTCCAGAAACAGTCTCCCGCAGA 324
QY 137 GCGCGCTGCGAGGCCCAACAGAGTGTGAGAACAGAGTGTGGAGGAGCGCTCCAGGC 196
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 325 GCGCGCTGTCAACCCACACACAGAGTGTGAGAGCCAGGCTGTGGAGGAGCTTCAGGT 384
QY 197 ACTGCCAGTCCGACACAAACCTGCAAAATCCATTAGAG 235
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 385 ACCTCGTACTGTACAACAACATCTGTAAATCCAACCGAG 423
```

Search completed: April 15, 2003, 18:45:43
Job time : 52.2209 secs


```
QY 1 CGAGGTACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAAGCCGAGCTCAA 60
    |||
Db 239 CGAGGTACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAAGCCGAGCTCAA 238
QY 61 AGATGAAGTTGGGAAGGTTAACAACCACTGCTGCCCTGCAAGGCGGACTT--CCAG 118
    |||
Db 299 AGATGAAGTTGGGAAGGTTAACAACCACTGCTGCCCTGCAAGGCGGACTTCCAGA 358
QY 119 AATACCTCTCCCGCGCGCTGCGAGCGCGCTGCGAGCGGTTGTAAGCAAGCTCTG 178
    |||
Db 359 AGTACCTCTCCCGCGCGCTGCGAGCGCGCTGCGAGCGGTTGTAAGCAAGGCTCTG 418
QY 179 GTGAGGCGAGCTCCAGGCTGCGAGCTGCCAGTCCAGCAACCTGCAAAATCCATTAGAGCCA 238
    |||
Db 419 GTGAGGCGAGCTCCAGGCTGCGAGCTGCCAGTCCAGCAACCTGCAAAATCCATTAGAGCCA 478
QY 239 CTG 241
    |||
Db 479 CTG 481
```

```
RESULT 11
US-09-907-372-16
; Sequence 16, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 702245091H1
US-09-907-372-16
```

```
Query Match 26.5%; Score 171.8; DB 10; Length 371;
Best Local Similarity 81.0%; Pred. No. 8.2e-44;
Matches 200; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
QY 402 GGTCTCCCGCAGCCCGAGTTTGGAGGCGAGGGTGCCGCAACAGCAGATCTCTGAGCC 461
    |||
Db 4 GACTCCCGCAGCCCGCAGTTGTGGAGAGAGATGCTACAGCAGATGCTCTGAGCC 63
QY 462 TGACCAAGGAGCGCAGTTGGAACCCGCGGAGCAGAGCCAGGTGGCCACGGTACCAATG 521
    |||
Db 64 AGGCAAGAGAGCTGAGCTGAGCTCCAGAACAGAGCCAGGTGGCCACGGTACCAATG 123
QY 522 GCATTCATGTCACCGGGGCTATGACTATGCTATGCAACATCTAATCTACAAATGGAC 581
    |||
Db 124 GCATTCATGTCACCGGGGCTATGACTATGCTATGCAACATCTAATCTACAAATGGAC 183
QY 582 CAGTACTGGGGGAGCACCAGGCTCTGAGAGACCTCCAGGTACCCCGCAACTCTCAATACC 641
    |||
Db 184 CAGTACTGGGGGAGCACCAGGCTCTGAGAGACCTCCAGGTACCCCGCAACTCTCAATACC 243
QY 642 CCAATTC 648
    |||
Db 244 CCAATTC 250
```

```
RESULT 12
US-09-907-372-5
; Sequence 5, Application US/09907372
; Patent No. US20020068242A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 4048821H1
US-09-907-372-5
```

```
Query Match 25.3%; Score 164.4; DB 10; Length 436;
Best Local Similarity 99.4%; Pred. No. 1.7e-41;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CGAGGTACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAAGCCGAGCTCAA 60
    |||
Db 136 CGAGGTACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAAGCCGAGCTCAA 195
QY 61 AGATGAAGTTGGGAAGGTTAACAACCACTGCTGCCCTGCAAGGCGGACTTCCAGAA 120
    |||
Db 196 AGATGAAGTTGGGAAGGTTAACAACCACTGCTGCCCTGCAAGGCGGACTTCCAGAA 255
QY 121 TACCTCTCCCGCAGCGCGCTGCGAGCGCGCCAGCCAGCAGGTGTGAG 166
    |||
Db 256 TACCTCTCCCGCAGCGCGCTGCGAGCGCGCCAGCCAGCAGGTGTGAG 301
```

```
RESULT 13
US-09-907-372-14
; Sequence 14, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 14
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 702152066H1
US-09-907-372-14
```

```
Query Match 21.7%; Score 140.6; DB 10; Length 548;
Best Local Similarity 72.5%; Pred. No. 4.4e-34;
Matches 182; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
```

```
QY 17 GAGCTACTTCTGACTGCGCGCTGCGAGCTGAAGCCGAGCTCAAAAGATGAAGTTGGGAG 76
    |||
Db 272 GAGCGGCTTGTACTCTGCGCGCTGCGAGCTGAAGCCGAGCTCAAAAGATGAAGTTATGAT 331
QY 77 GGTAAACACCACTGCTGCCCTGCAAGGCGAGGCACTTCAGAAATACCTCTCCCGCAGC 136
    |||
Db 332 ACTGAAGTCAACTGTGTGCTCCCTGTAAGCCAGGACACTTCCAGAAACGTCTCCCGCAGA 391
QY 137 GCCCGTCCAGCCCGCAGCAGGTGTGAGAACCAAGAGTGTGGGAGGAGCTCCAGGC 196
    |||
Db 392 GCCCGTCCAGCCCGCAGCAGGTGTGAGAACCAAGAGTGTGGGAGGAGCTTCAGGT 451
```


Query Match 81.8%; Score 531; DB 10; Length 2136;
 Best Local Similarity 85.7%; Pred. No. 3e-155;
 Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

```

QY 1 CGAGTGTACACACTGCGAGCTACTTCTGTGACTGCGCCGCTGCGGCACTGAGCCGAGCTCAA 60
    |||
Db 579 CGAGTGTACACACTGCGAGCTACTTCTGTGACTGCGCCGCTGCGGCACTGAGCCGAGCTCAA 638
    |||
QY 61 AGATGAAGTGGAGAGGTAAACCAACCACTGCGTCCCTGGAAGGCAAGGCACTTCCGAA 120
    |||
Db 639 AGATGAAGTGGAGAGGTAAACCAACCACTGCGTCCCTGGAAGGCAAGGCACTTCCGAA 698
    |||
QY 121 TACCTCTCCCGCAGCGCCGCTGCGCAGCCCAACAGGTGTGAGAACCAAGGTCTGTT 180
    |||
Db 699 TACCTCTCCCGCAGCGCCGCTGCGCAGCCCAACAGGTGTGAGAACCAAGGTCTGTT 758
    |||
QY 181 GAGAGGCACTCCAGGCACTGCGCAGCAACCACTGCAAAAAATCATTTAGAGCCACT 240
    |||
Db 759 GAGAGGCACTCCAGGCACTGCGCAGCAACCACTGCAAAAAATCATTTAGAGCCACT 818
    |||
QY 241 GCGCCGAGAGATGCA----- 256
    |||
Db 819 GCGCCGAGAGATGCAAGAACCATGCTGATGCGCGCTTCTGCTGCCACTGCGCTTCTT 878
    |||
QY 257 ----- 256
    |||
Db 879 TCTGCTCTTGCCACCGCTTCTCTGCTGATCTGGAAGGCAACCGCTTCTCTGCAAGAA 938
    |||
QY 257 -----GATTCGCTGCTCAAGAGGCGTCCGCAAGGAGAGGCAACCGCTTCTGCTGAG 312
    |||
Db 939 ACTGGGATTCCTGCTCAAGAGGCGTCCGCAAGGAGAGGCAACCGCTTCTGCTGAG 998
    |||
QY 313 CTGGGACCTCTCCAGAGGCGCATTCATCTCCCTGACTTGGTATGAGCAGCTATGCCAT 372
    |||
Db 999 CTGGGACCTCTCCAGAGGCGCATTCATCTCCCTGACTTGGTATGAGCAGCTATGCCAT 1058
    |||
QY 373 TTCTGAGATGTTTCCCAATATCAGTGGGCTCCCGCAGCGCCAGTTTGGAGGAGG 432
    |||
Db 1059 TTCTGAGATGTTTCCCAATATCAGTGGGCTCCCGCAGCGCCAGTTTGGAGGAGG 1118
    |||
QY 433 GGTGCCCCAACAGAGAGTCTCTGACCTGACAGGAGGCGGAGTTTGGAGGAGG 492
    |||
Db 1119 GGTGCCCCAACAGAGAGTCTCTGACCTGACAGGAGGCGGAGTTTGGAGGAGG 1178
    |||
QY 493 GCAGAGCCAGGTGGCCAGGTTACCATGATTCATGACCGGCGGTCTATGACTAT 552
    |||
Db 1179 GCAGAGCCAGGTGGCCAGGTTACCATGATTCATGACCGGCGGTCTATGACTAT 1238
    |||
QY 553 CACTGGCAACATCTACATCTACATGAGCAGTACTGGGGGAGCACCGGCTCTGAGA 612
    |||
Db 1239 CACTGGCAACATCTACATCTACATGAGCAGTACTGGGGGAGCACCGGCTCTGAGA 1298
    |||
QY 613 CTTCCAGCTACCCCGCAACCTCCATACCCCATTTCCC 649
    |||
Db 1299 CTTCCAGCTACCCCGCAACCTCCATACCCCATTTCCC 1335
    |||

```

RESULT 6
 US-09-907-372-8
 ; Sequence 8, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti G.
 ; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
 ; FILE REFERENCE: PC-0050 US
 ; CURRENT APPLICATION NUMBER: US/09/907,372
 ; CURRENT FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 8
 ; LENGTH: 574

TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 OTHER INFORMATION: Incyte ID No. US20020068242A1 8234468H1
 US-09-907-372-8

Query Match 72.8%; Score 472.4; DB 10; Length 574;
 Best Local Similarity 99.8%; Pred. No. 3.6e-137;
 Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 176 CTGGTGGAGGCACTCCAGGCACTGCGCAGTCCGACAACTGCAAAATTCATTAGAG 225
    |||
Db 1 CTGGTGGAGGCACTCCAGGCACTGCGCAGTCCGACAACTGCAAAATTCATTAGAG 60
    |||
QY 236 CCAGTGGCCCGCAGAGATGAGATGCTGCTCAAGAGGCGTCCGAGGAGAGGAGCC 295
    |||
Db 61 CCAGTGGCCCGCAGAGATGAGATGCTGCTCAAGAGGCGTCCGAGGAGAGGAGCC 120
    |||
QY 296 AATCTGTACCTGGAAGCTGGAGCCTCCGAAAGGCCCATCATCTTCCCTGACTTGTA 355
    |||
Db 121 AATCTGTACCTGGAAGCTGGAGCCTCCGAAAGGCCCATCATCTTCCCTGACTTGTA 180
    |||
QY 356 CAGCCACTGCTACCATTTCTGAGATGTTTCCCAATATCAGTGGGCTCCCGCAGCC 415
    |||
Db 181 CAGCCACTGCTACCATTTCTGAGATGTTTCCCAATATCAGTGGGCTCCCGCAGCC 240
    |||
QY 416 CCAGTTTGGAGGAGGAGGCGCCGCAACAGAGATCTGCGACCTGAGCAGGAGCCG 475
    |||
Db 241 CCAGTTTGGAGGAGGAGGCGCCGCAACAGAGATCTGCGACCTGAGCAGGAGCCG 300
    |||
QY 476 CAGTTGGAACCCGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 535
    |||
Db 301 CAGTTGGAACCCGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
    |||
QY 536 GCGGCGCTTATGACTATACATGCGCAACATCTACATCTACATGAGCAGTATGAGG 595
    |||
Db 361 GCGGCGCTTATGACTATACATGCGCAACATCTACATCTACATGAGCAGTATGAGG 420
    |||
QY 596 CCACCGGCTCTGAGAGCTCCAGTACCCCGCAACCTCCATACCCCATTTCCC 649
    |||
Db 421 CCACCGGCTCTGAGAGCTCCAGTACCCCGCAACCTCCATACCCCATTTCCC 474
    |||

```

RESULT 7
 US-09-907-372-9/c
 ; Sequence 9, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti G.
 ; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
 ; FILE REFERENCE: PC-0050 US
 ; CURRENT APPLICATION NUMBER: US/09/907,372
 ; CURRENT FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 9
 ; LENGTH: 425
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. US20020068242A1 7716340H1
 US-09-907-372-9

Query Match 65.5%; Score 425; DB 10; Length 425;
 Best Local Similarity 100.0%; Pred. No. 1.8e-122;
 Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 225 ATCCATTAGAGCCACTGCCCCAGAGATGTCAGATGCTGCTTAAGAGGCGTCCGAGG 284
    |||
Db 425 ATCCATTAGAGCCACTGCCCCAGAGATGTCAGATGCTGCTTAAGAGGCGTCCGAGG 366
    |||

```

```

Db 759 GGAGGACAGCTCCAGGACATGCGCCAGTCCGACACCACTGCAAAAATCATATAGGCCACT 818
QY 241 GCGCCGACAGATGCA----- 256
Db 819 GCGCCGACAGATGCAAGAACCAATGCTGAGTGGCGGTTCTGCTGCCACTGGCTTCTT 878
QY 257 ----- 256
Db 879 TCTGCTCTTGGCAGCGTCTTCTGCTGATCTGGAAGACCACTTCTCTGCGAGAA 938
QY 257 -----GATCGCTGCTCAAGAGGCGTCCGACAGGAGAGGAGCAATCCTGTAGCTGGAG 312
Db 939 ACTGGGATCGCTGCTCAAGAGGCGTCCGACAGGAGAGGAGCAATCCTGTAGCTGGAG 998
QY 313 CTGGAGAGCTCCGAGAGGCGCAATCATCTTCCCTGACTTGGTATAGGCACTGCTACCAT 372
Db 999 CTGGAGAGCTCCGAGAGGCGCAATCATCTTCCCTGACTTGGTATAGGCACTGCTACCAT 1058
QY 373 TTCTGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCACTTTTGGAGGAGG 432
Db 1059 TTCTGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCACTTTTGGAGGAGG 1118
QY 433 GGTGCGGCAACAGCAGAGTCTCTGACCTGACCAAGGAGCCGCAATTGGAAACCCGGGGA 492
Db 1119 GGTGCGGCAACAGCAGAGTCTCTGACCTGACCAAGGAGCCGCAATTGGAAACCCGGGGA 1178
QY 493 GCAGAGCAGAGTGGCGGCGGATGCAATGCAATTCATGACCGGGGGCTATGACTAT 552
Db 1179 GCAGAGCAGAGTGGCGGCGGATGCAATGCAATTCATGACCGGGGGCTATGACTAT 1238
QY 553 CACTGGCAACATCTATCATATGACAGCAGTACTGAGGGGAGCACCAGGGTCTTGAGAG 612
Db 1239 CACTGGCAACATCTATCATATGACAGCAGTACTGAGGGGAGCACCAGGGTCTTGAGAG 1298
QY 613 CCTCCAGCTACCGCCGAGACCTTCATACCCCATTTCC 649
Db 1299 CCTCCAGCTACCGCCGAGACCTTCATACCCCATTTCC 1335

RESULT 4
US-09-962-436-262
; Sequence 262, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 262
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-262

Query Match 81.8%; Score 531; DB 10; Length 2136;
Best Local Similarity 85.7%; Pred. No. 3e-155;
Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

```

```

Db 639 AGATGAAGTTGGAGAGGTATACAAACCACTGCTCCCTGCAAGGCGAGGCACTTCCAGAA 698
QY 121 TACCTCTCCCGCAGGCGCCGCTGCGCAGGCCACACAGAGTGTAGAACCAAGTCTGGT 180
Db 699 TACCTCTCCCGCAGGCGCCGCTGCGCAGGCCACACAGAGTGTAGAACCAAGTCTGGT 758
QY 181 GGAGGCAAGCTCCAGGCACTGCGCAATGTCGACACACACCTTCGAAAAATCATATTAGAGCACT 240
Db 759 GGAGGCAAGCTCCAGGCACTGCGCAATGTCGACACACACCTTCGAAAAATCATATTAGAGCACT 818
QY 241 GCGCCGACAGATGCA----- 256
Db 819 GCGCCGACAGATGCAAGAACCAATGCTGATGCTGGCGGTTCTGCTGCCACTGGCTTCTT 878
QY 257 ----- 256
Db 879 TCTGCTCTTGGCAGCGTCTTCTGCTGATCTGGAAGACCACTTCTCTGCGAGAA 938
QY 257 -----GATCGCTGCTCAAGAGGCGTCCGACAGGAGAGGAGCAATCCTGTAGCTGGAG 312
Db 939 ACTGGGATCGCTGCTCAAGAGGCGTCCGACAGGAGAGGAGCAATCCTGTAGCTGGAG 998
QY 313 CTGGAGAGCTCCGAGAGGCGCAATCATCTTCCCTGACTTGGTATAGGCACTGCTACCAT 372
Db 999 CTGGAGAGCTCCGAGAGGCGCAATCATCTTCCCTGACTTGGTATAGGCACTGCTACCAT 1058
QY 373 TTCTGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCACTTTTGGAGGAGG 432
Db 1059 TTCTGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCACTTTTGGAGGAGG 1118
QY 433 GGTGCGGCAACAGCAGAGTCTCTGAGCTGACCAAGGAGCCGCAATTGGAAACCCGGGGA 492
Db 1119 GGTGCGGCAACAGCAGAGTCTCTGAGCTGACCAAGGAGCCGCAATTGGAAACCCGGGGA 1178
QY 493 GCAGAGCAGAGTGGCGGCGGATGCAATGCAATTCATGACCGGGGGCTATGACTAT 552
Db 1179 GCAGAGCAGAGTGGCGGCGGATGCAATGCAATTCATGACCGGGGGCTATGACTAT 1238
QY 553 CACTGGCAACATCTATCATATGACAGCAGTACTGAGGGGAGCACCAGGGTCTTGAGAG 612
Db 1239 CACTGGCAACATCTATCATATGACAGCAGTACTGAGGGGAGCACCAGGGTCTTGAGAG 1298
QY 613 CCTCCAGCTACCGCCGAGACCTTCATACCCCATTTCC 649
Db 1299 CCTCCAGCTACCGCCGAGACCTTCATACCCCATTTCC 1335

RESULT 5
US-09-880-107-2135
; Sequence 2135, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2135
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 I04270
US-09-880-107-2135

```

Qy	241	GGCCCCAGAGATGTGACAGATCGCTGCTCAAGAGCGCTCCGACAGGAGAGAGACCAATCC	300
Db	869	GGCCCCAGAGATGTGACAGATCGCTGCTCAAGAGCGCTCCGACAGGAGAGAGAGACCAATCC	928
Qy	301	TGTAGCTGGAAGCTGGGAGGCTCCGAAGGGCCATCATCTTCCGTGATTTGGTACATCC	360
Db	929	TGTAGCTGGAAGCTGGGAGGCTCCGAAGGGCCATCATCTTCCGTGATTTGGTACATCC	988
Qy	361	ACTGCTACCACTTTCTGGAGATGTTTCCCAAGTATCCACTTGGGCTCCCGCAGCCCACT	420
Db	989	ACTGCTACCACTTTCTGGAGATGTTTCCCAAGTATCCACTTGGGCTCCCGCAGCCCACT	1048
Qy	421	TTTGGAGGCGAGGGTGCCCGCAACAGCAGATCTCTGTGACCTGTACCAAGGAGCCCACTT	480
Db	1049	TTTGGAGGCGAGGGTGCCCGCAACAGCAGATCTCTGTGACCTGTACCAAGGAGCCCACTT	1108
Qy	481	GGAACCCGGGGAGCAGAGCCAGGAGGGGCCAGGTACTACATGGATTCATGTACCGGGGG	540
Db	1109	GGAACCCGGGGAGCAGAGCCAGGAGGGGCCAGGTACTACATGGATTCATGTACCGGGGG	1168
Qy	541	GTCATAGCATATCACTGTGCAACATCTACATCTACATGAGACCAAGTACTGGGGGAGCAC	600
Db	1169	GTCATAGCATATCACTGTGCAACATCTACATCTACATGAGACCAAGTACTGGGGGAGCAC	1228
Qy	601	GGGTCTGTGAGACCTCCCACTACACCCCGAAGCTCCATACCCCATTTCCC	649
Db	1229	GGGTCTGTGAGACCTCCCACTACACCCCGAAGCTCCATACCCCATTTCCC	1277

```

RESULT 2
US-09-907-372-7/c
Sequence 7, Application US/09907372
Patent No. US20020068242A1
GENERAL INFORMATION:
APPLICANT: Lal, Preeti G.
APPLICANT: Warren, Bridget A.
TITLE OF INVENTION: TUN RECEPTOR 2 RELATED PROTEIN VARIANT
FILE REFERENCE: PC-0050 US
CURRENT APPLICATION NUMBER: US/09/907,372
CURRENT FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 651
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID NO. US20020068242A1 7716364H
US-09-907-372-7

```

Query Match	92.6%;	Score 600.8;	DB 10;	Length 651;
Best Local Similarity	99.2%;	Pred. No. 4.7e-177;		
Matches 635;	Conservative	0;	Mismatches 2;	Indels 3; Gaps 3

QY	13	CTGCAGGCTACTTTTGTACTCTCCCGCCGGGACATCGAAGCCGAGTCTCAAAGTGAAGTTGG	72
Db	640	CTGCAGGCTACTTTTGTACTCTCCCGCCGGGACATCTGTACCCGAGTCTCAAAGTGAAGTTGG	581
QY	73	GAAGGGTAAACAACCACTGCGTCCCTGTGCAGGCGAGGCGACTTCAGAAATACCTCTCC	132
Db	580	GAAGGGTAAACAACCACTGCGTCCCTGTGCAGGCGAGGCGACTTCAGAAATACCTCTCC	521
QY	133	CAGGCCCC - GCTGCCAGGCCCCACACAGGTTGTGAAGAACCAAGTCTGTGGAGGCACTC	191
Db	520	CAGGCCCTCGTGGCCAGGCCCCACACAGGTTGTGAAGAACCAAGTCTGTGGAGGCACTC	461
QY	192	CAGGCACATGCCCACATGTCGACACAAACCTGCAAAAATCCATTGTAGACCACTGCCCCAGAGA	251
Db	460	CAGGCACATGCCCACATGTCGACACAAACCTGCAAAAATCCATTGTAGACCACTGCCCCAGAGA	401
QY	252	TGTCAGATTCGCTGCTCAAGAGGGGTCCGCAAGGAGAGAGGAGCCCAATCTGTAGCTGAA	311

Db	400	TEICAGAGATCCGCTGCTCAAGAGGGGTCCGCAGAGGAGAGGCCCAATCCTGTAGCTGGAA	34.1
Qy	312	GCTGGGAGCCCTCCGAAGGCCCATCCATACCTTCCTGACTTGGTACAGCCCACTGCTACCA	37.1
Db	340	GCTGGAGACCTCTCGAAGGCCCATCCATACCTTCCTGACTTGGTACAGCCCACTGCTACCA	28.1
Qy	372	TTTCTGAGAGATGTTTCCCACTATTCACACTGGGCTCCCCGACGCCAGTTTGGAGGCAG	43.1
Db	280	TTTCTGAGAGATGTTTCCCACTATTCACACTGGGCTCCCCGACGCCAGTTTGGAGGCAG	22.1
Qy	432	GGGTCCGCACACAGAGAGGTCTCTGGACCTGACACAGGAGGCCGACTTTGGAAACCCGGGG	49.1
Db	220	GGGTCCGCACACAGAGAGGTCTCTGGACCTGACACAGGAGGCCGACTTTGGAAACCCGGGG	16.1
Qy	492	AGCAGAGCCAGGTGGCCCGACGGTACCAATGGCATTCATGTCACGGCGGGTCTATGACTA	55.1
Db	160	AGCAGAGCCAGGTGGCCCGACGGTACCAATGGCATTCATGTCACGGCGGGTCTATGACTA	10.0
Qy	552	TCACGTGGCAACATCTACATCTACAAATGGACACAGTATGGGGGGACACACCGGGTCTGGAG	61.1
Db	100	TCACGTGGCAACATCTACATCTACAAATGGACACAGTATGGGGGGACACACCGGGTCTGGAG	4.1
Qy	612	ACCTCCG--AGCTACCCCGGAACGCTCC--ATACCACATTCGCC	64.9
Db	40	ACCTCCCTAGCTACCCCGGAACGCTCTCTATACCCCATTTCCC	1

```

RESULT 3
US-09-954-531-1348
: Sequence 1348, Application US/09954531
: Patent No. US20020165180A1
: GENERAL INFORMATION:
: APPLICANT: Weaver, Zoe
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
: TITLE OF INVENTION: Gene Sets
: FILE REFERENCE: 689290-77
: CURRENT APPLICATION NUMBER: US/09/954, 531
: CURRENT FILING DATE: 2002-05-02
: PRIOR APPLICATION NUMBER: US/60/233, 133
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234, 009
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234, 034
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234, 509
: PRIOR FILING DATE: 2000-09-22
: PRIOR APPLICATION NUMBER: US/60/234, 567
: PRIOR FILING DATE: 2000-09-22
: NUMBER OF SEQ ID NOS: 1392
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1348
: LENGTH: 2136
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-954-531-1348

```

Query Match	81.8%;	Score 531;	DB 9;	Length 2136;
Best Local Similarity	85.7%;	Pred. No. 3e-155;		
Matches 649;	Conservative 0;	Mismatches 0;	Indels 108;	Gaps 1;

QY	1	CGAGTATACACACTGGGAGCTACTTCTGACTGCCCGCTGGCAGTGAAGCCGAGCTCAA	60
Db	579	CGAGTGTACACACTGGGAGCTACTTCTGACTGCCCGCTGGCAGTGAAGCCGAGCTCAA	638
QY	61	AGATGAAGTTGGGAAGGTATACAAACCACATGCGTCCCTCAAGGCGGGGACCTTCCAGAA	120
Db	639	AGATGAAGTTGGGAAGGTATACAAACCACATGCGTCCCTCAAGGCGGGGACCTTCCAGAA	698
QY	121	TACCTCTCTCCCGCAGGCGCGCTGCCAGGCCCAACACAGTGTGAAGAACCAAGTGTGCT	180
Db	699	TACCTCTCTCCCGCAGGCGCGCTGCCAGGCCCAACACAGTGTGAAGAACCAAGTGTGCT	758
QY	181	GGAGGAGCTCCAGGCACACTGATCCACACAACTCTCAAAAATCCATTGAGAGCCACT	240

QY	61	AGATGAAGTTGGGAAGGGTAACAAACCACTGCTGCTCCCTGCAGAGCGAGGGCACTTCCAGAA	120
QY	1	CGAGGTTCACACACTGCGAGAGCTACTTTTGTGATGTGCCGCGCGTGGACATGAACCCAGCTCAAA	60
QY	629	CGAGGTACACACTGCCAGCTACTTTGTGACTGTGCCGCGCTGGACACTGAACCCAGAGCTCAA	688
Db	61	AGATGAAGTTGGGAAGGGTAACAAACCACTGCTGCTCCCTGCAGAGCGAGGGCACTTCCAGAA	120
Db	689	AGATGAAGTTGGGAAGGGTAACAAACCACTGCTGCTCCCTGCAGAGCGAGGGCACTTCCAGAA	748
QY	121	TACCTCCCTCCCCAGCGCGCGCTGCCAGCCCCCAACCAAGGTGTGAGAAACCAAGGTCGT	180
QY	749	TACCTCCCTCCCCAGCGCGCGCTGCCAGCCCCCAACCAAGGTGTGAGAAACCAAGGTCGT	808
Db	181	GGAGGCACCTCCAGGCACTGCGCCAGTCCGACACAACTGCAAAAATCATTTAGAGCCACT	240
QY	809	GGAGGCACCTCCAGGCACTGCGCCAGTCCGACACAACTGCAAAAATCATTTAGAGCCACT	868

QY	80	AACACACCTCTGTCCTCCCTGGAAGCAGGAGACTTCCAGAAATACCTCTCCCCCAAGGCC	139
Db	756	AACACGAGTGTCCAGCCGTGCCCCCAGGACCTTCTCAGCCACGACGCTCCAGCTCAGAG	815
QY	140	CGATGCCAGCCCCACACACAGGTGTGAATCCAAAGGTGTGGGAGGCGAGGTCCAGGCACCT	199
Db	816	CAGTGCACGCCCCCACCAGACTGACGAGCGCCCTGGCCTCCTCAATGTGCGAGGCTCT	875
QY	200	GCCGAGTCCGACACACCTGCA	221
Db	876	TCTCTCCATGACACCTCTGTGCA	897

RESULT 15
US-09-286

; Sequence 19, Application US/09286529

; Patent NO. 629/36/
; GENERAL INFORMATION:

```

; APPLICANT: Catherine Tribouley
; TITLE OF INVENTION: NEW MEMBERS OF TNF AND TNFR FAMILIES

```

FILE REFERENCE: 1408.003/200130.439C1
CURRENT APPLICATION NUMBER: US/09/286.529

CURRENT FILING DATE: 1999-04-05

```
; NUMBER OF SEQ ID NOS: 25
;
; SOFTWARE: FastSeq for Windows Version 3.0
```

```

; SEQ ID NO 19
;
; LENGTH: 1859

```

TYPE: DNA
ORGANISM: Homo sapien

US-09-286-529-19

Query Match	5.6%	Score	36.4	DB	4	Length	1859
Best Local Similarity	53.5%	Pred. No.	0.54				
Matches	76	Conservative	0	Mismatches	66	Indels	0
						Gaps	0

QY	80	AACAAACACTCTGTCCTCCCTGCAAGGACGGGCACTTCCAGAAATACCTCTCTCCCAAGGCC	139
Db	516	AACAGCAGTGTCCAGCCGTGCCCCCAGGCACTTCTCAGCAGACAGCTCCAGCTCAGAG	575
QY	140	CGATGCCAGCCCCCAACCAGGTGTGAGAACCAAGGTGTGGTGGAGGAGCACTCCAGGCACCT	199
Db	576	CAGTGCAGGCCCAACCGCAATGTGACAGGCCCTTGAGCCTTCAATGTGCCAGGCTCT	635
QY	200	GCCCACTCCGACACAACTGCA	221
Db	636	TCTCTCCATGACACCTGTGCA	657

Search completed: April 15, 2003, 16:28:35
Job time : 29.9124 secs

SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1004
TYPE: DNA
ORGANISM: Homo sapiens
US-08-114-944D-1

Query Match
Best Local Similarity 56.1%; Pred. No. 0.074;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 79 TAAACACACTGCGTCCCTGCAAGCAGGCGACTTCAGAAATCCCTCCAGGCG 138
DB 464 TGATACACATCTGCGAGCGCTCCAGTGGCTTCTTCCAAATGTGATCTGCTTGA 523
QY 139 CCGCTGCGAGCCCGACACACAGAGTGTGAGAACCAAGTGTGTGAGGCGAGCTCCAGGCAC 198
DB 524 AAAATGTACACCTGTGAGAACCTGTGAGAACCAAGACCTGTGTGCAACAGCAGGCGAC 583
QY 199 TGCCGAGTCCGA 210
DB 584 AACACAGACTGA 595

RESULT 12
US-09-286-529-7
Sequence 7, Application US/09286529
Patent No. 6297367
GENERAL INFORMATION:
APPLICANT: Catherine Tribouley
TITLE OF INVENTION: NEW MEMBERS OF TNF AND TNFR FAMILIES
FILE REFERENCE: 1408.003/200130.439C1
CURRENT APPLICATION NUMBER: US/09/286,529
CURRENT FILING DATE: 1999-04-05
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 459
TYPE: DNA
ORGANISM: human
US-09-286-529-7

Query Match
Best Local Similarity 53.5%; Pred. No. 0.35;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 80 AACACACACTGCGTCCCTGCAAGCAGGCGACTTCAGAAATCCCTCCAGGCGC 139
DB 187 AACACGAGTGTGAGAACCAAGTGTGTGAGGCGAGCTCCAGGCGACT 246
QY 140 CGCTGCGAGCCCGACACACAGAGTGTGAGAACCAAGTGTGTGAGGCGAGCTCCAGGCGACT 199
DB 247 CAGTGGCAGGCCCGACCGCACTTCAGGCGCTGCGCTGAGCCTCAATGTGCGAGGCTCT 306
QY 200 GCCCAGTCCGACACAACCTGCA 221
DB 307 TCCTCCCATGACACCTGTGCA 328

RESULT 13
US-08-794-796-1
Sequence 1, Application US/08794796
Patent No. 5885800
GENERAL INFORMATION:
APPLICANT: Emery, John
APPLICANT: Tan, KB
APPLICANT: Truneh, Alem
APPLICANT: Young, Peter
TITLE OF INVENTION: Tumor Necrosis Related Receptor,
TITLE OF INVENTION: TR4
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/794,796
FILING DATE: 04-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: GH50000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-4026
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1164 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-794-796-1

Query Match
Best Local Similarity 53.5%; Pred. No. 0.47;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 80 AACACACACTGCGTCCCTGCAAGCAGGCGACTTCAGAAATCCCTCCAGGCGC 139
DB 549 AACACGAGTGTGAGAACCAAGTGTGTGAGGCGAGCTCCAGGCGACT 608
QY 140 CGCTGCGAGCCCGACACACAGAGTGTGAGAACCAAGTGTGTGAGGCGAGCTCCAGGCGACT 199
DB 609 CAGTGGCAGGCCCGACCGCACTTCAGGCGCTGCGCTGAGCCTCAATGTGCGAGGCTCT 668
QY 200 GCCCAGTCCGACACAACCTGCA 221
DB 669 TCCTCCCATGACACCTGTGCA 690

RESULT 14
US-09-286-529-18
Sequence 18, Application US/09286529
Patent No. 6297367
GENERAL INFORMATION:
APPLICANT: Catherine Tribouley
TITLE OF INVENTION: NEW MEMBERS OF TNF AND TNFR FAMILIES
FILE REFERENCE: 1408.003/200130.439C1
CURRENT APPLICATION NUMBER: US/09/286,529
CURRENT FILING DATE: 1999-04-05
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 18
LENGTH: 1347
TYPE: DNA
ORGANISM: Homo sapien
US-09-286-529-18

Query Match
Best Local Similarity 53.5%; Pred. No. 0.49;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

RESULT 10
US-09-041-886-26
; Sequence 26, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:

RESULT 11
US-08-114-944D-1
Sequence 1, Application US/08114944D
Patent No. 6376459
GENERAL INFORMATION:
APPLICANT: Aruffo, Alejandro J
APPLICANT: Ledbetter, Jeffrey A
APPLICANT: Stamenkovic, Ivan
APPLICANT: No. 637645911e, Randolph
TITLE OF INVENTION: THE CD40CR RECEPTOR AND LIGANDS THEREFOR
FILE REFERENCE: 5624-232-999
CURRENT APPLICATION NUMBER: US/08/114,944D
CURRENT FILING DATE: 1993-08-31
PRIOR APPLICATION NUMBER: 07/835,799
PRIOR FILING DATE: 1992-02-14
NUMBER OF SEQ ID NOS: 3

QY	199	TGCCAGTCCGA	210
Db	479	AAACAAGACTGA	490

```

RESULT 7
US-08-770-981-4
: Sequence 4, Application US/08770981
: Patent No. 6391637
: General Information:
: APPLICANT: ARMITAGE, RICHARD
: APPLICANT: FANSHAW, WILLIAM
: APPLICANT: SPRIGGS, MELANIE
: APPLICANT: SRINIVASAN, SUBHASHINI
: APPLICANT: GIBSON, MARLOU
: TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: IMMUNEX CORPORATION
: STREET: 51 UNIVERSITY STREET
: CITY: SEATTLE
: STATE: WASHINGTON
: COUNTRY: USA
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: Apple Macintosh
: OPERATING SYSTEM: Apple Operating System 7.1
: SOFTWARE: Microsoft Word for Apple, version 5.1a
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/770,981
: FILING DATE: 20-DEC-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/477,733
: FILING DATE: 07-JUN-1995
: APPLICATION NUMBER: 08/249,189
: FILING DATE: May 24, 1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/969,703
: FILING DATE: October 23, 1992
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/805,723
: FILING DATE: December 5, 1991
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/783,707
: FILING DATE: October 25, 1991
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Perkins, Patricia A.
: REGISTRATION NUMBER: 34,693
: REFERENCE/DOCKET NUMBER: 2802-D
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 2065870430
: TELEFAX: 2065870606
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 519 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: HUMAN
: IMMEDIATE SOURCE:
: CLONE: CD40 EXTRACELLULAR REGION
: IS-08-770-981-4

```

Query Match	6.0%	Score 39.2	DB 4	Length 519
Best Local Similarity	56.1%	Pred. No. 0.06		
Matches 74	Conservative 0	Mismatches 58	Indels 0	Gaps 0
QY	79	TAACACCACTGCGCTGCGCCCTGCAGACGAGGAGGACACTTCCAGAAATACCTCTCCCTCCCGACGC	138	
Db	359	TGATATACATCTGGCGAGCGCCTGCCAGTGGCGTCTTCTTCACATGTGTATCTGCTTTGCA	418	
QY	139	CCGCTGCCAGCCCCACACCCAGAGGTGTAGAAACCAAGGTCTGTTGGAGGCAAGCTCCAGGCAC	198	
Db	419	AAAATGTCAACCCCTTGGACAAAGCTGTGACACCAAAAAGACCTGTTGTGCAACAGGACGCGAC	478	
QY	199	TGCCAGTTCGCA	210	
Db	479	AAACAGACTGA	490	

```

RESULT 8
US-09-399-106-4
: Sequence 4, Application US/09399106
: Patent No. 6410711
:
: GENERAL INFORMATION:
: APPLICANT: ARMITAGE, RICHARD
: APPLICANT: FANSLAW, WILLIAM
: APPLICANT: SPRIGGS, MELANIE
: APPLICANT: SRINIVASAN, SUBHASHINI
: APPLICANT: GIBSON, MARYLOU
: APPLICANT: MORRIS, ARVIA E.
: APPLICANT: MCGREW, JEFFERY
: TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
: NUMBER OF SEQUENCES: 26
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: IMMUNEX CORPORATION
: STREET: 51 UNIVERSITY STREET
: CITY: SEATTLE
: STATE: WASHINGTON
: COUNTRY: USA
: ZIP: 98101
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: Apple Macintosh
: OPERATING SYSTEM: Apple Operating System 7.5.5
: SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/399,106
: FILING DATE:
:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/477,733
: FILING DATE:
:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/969,703
: FILING DATE: October 23, 1992
:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/805,723
: FILING DATE: December 5, 1991
:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/783,707
: FILING DATE: October 25, 1991
:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Perkins, Patricia A.
: REGISTRATION NUMBER: 34,693
: REFERENCE/DOCKET NUMBER: 2802-D
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 2065870430
: TELEFAX: 2065870606
:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:

```

1 TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
2 NUMBER OF SEQUENCES: 24
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: IMMUNEX CORPORATION
5 STREET: 51 UNIVERSITY STREET
6 CITY: SEATTLE
7 STATE: WASHINGTON
8 COUNTRY: USA
9 ZIP: 98101
10
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: Apple Macintosh
14 OPERATING SYSTEM: Apple Operating System 7.1
15 SOFTWARE: Microsoft Word for Apple, version 5.1a
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/08/769,819
18 FILING DATE: 19-DEC-1996
19 CLASSIFICATION: 424
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER: 08/484,624
22 FILING DATE: 07-JUN-1995
23 APPLICATION NUMBER: 08/249,189
24 FILING DATE: May 24, 1994
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: 07/969,703
27 FILING DATE: October 23, 1992
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: 07/805,723
30 FILING DATE: December 5, 1991
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER: 07/783,707
33 FILING DATE: October 25, 1991
34 ATTORNEY/AGENT INFORMATION:
35 NAME: Perkins, Patricia A.
36 REGISTRATION NUMBER: 34,693
37 REFERENCE/DOCKET NUMBER: 2802-E
38 TELECOMMUNICATION INFORMATION:
39 TELEPHONE: 2065870430
40 TELEFAX: 2065870606
41 INFORMATION FOR SEQ ID NO: 4:
42 SEQUENCE CHARACTERISTICS:
43 LENGTH: 519 base pairs
44 TYPE: nucleic acid
45 STRANDEDNESS: single
46 TOPOLOGY: linear
47 MOLECULE TYPE: cDNA
48 HYPOTHETICAL: NO
49 ANTI-SENSE: NO
50 ORIGINAL SOURCE:
51 ORGANISM: HUMAN
52 IMMEDIATE SOURCE:
53 CLONE: CD40 EXTRACELLULAR REGION
54 US-08-769-819-4

Query Match 6.0%; Score 39.2; DB 4; Length 519;
Best Local Similarity 56.1%; Pred. No. 0.06;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 79 TAACACCACTGCTCCCTGCAAGGAGGCACTTCAGATACCTCTCCCGCAGCGC 138
DB 359 TGATACATCTCGAGCGCTGCGCCAGCTGCTTCTCCATGTGTCATCTGCTTGA 418
QY 139 CCGCTGCAGCCCAACACAGGTGTGAGAACCAAGTGTGTGAGGAGAGTCCAGGAC 198
DB 419 AAATGTCACCTTGCAACAGCTGTGAGACCAAGAGACCTGTTGTGCAACAGGAGCGAC 478
QY 199 TGCCAGTCCGA 210
DB 479 AAACAAGACTGA 490

RESULT 6
US-08-770-974-4

1 Sequence 4, Application US/08770974
2 Patent No. 6290972
3 GENERAL INFORMATION:
4 APPLICANT: ARMITAGE, RICHARD
5 APPLICANT: FANSLON, WILLIAM
6 APPLICANT: SPRIGGS, MELANIE
7 APPLICANT: SRINIVASAN, SUBHASHINI
8 APPLICANT: GIBSON, MARYLOU
9 TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
10 NUMBER OF SEQUENCES: 24
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: IMMUNEX CORPORATION
13 STREET: 51 UNIVERSITY STREET
14 CITY: SEATTLE
15 STATE: WASHINGTON
16 COUNTRY: USA
17 ZIP: 98101
18
19 COMPUTER READABLE FORM:
20 MEDIUM TYPE: Floppy disk
21 COMPUTER: Apple Macintosh
22 OPERATING SYSTEM: Apple Operating System 7.1
23 SOFTWARE: Microsoft Word for Apple, version 5.1a
24 CURRENT APPLICATION DATA:
25 APPLICATION NUMBER: US/08/770,974
26 FILING DATE: 20-DEC-1996
27 CLASSIFICATION: 435
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: 08/477,733
30 FILING DATE: 02-AUG-1995
31 APPLICATION NUMBER: 08/249,189
32 FILING DATE: May 24, 1994
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: 07/969,703
35 FILING DATE: October 23, 1992
36 PRIOR APPLICATION DATA:
37 APPLICATION NUMBER: 07/805,723
38 FILING DATE: December 5, 1991
39 PRIOR APPLICATION DATA:
40 APPLICATION NUMBER: 07/783,707
41 FILING DATE: October 25, 1991
42 ATTORNEY/AGENT INFORMATION:
43 NAME: Perkins, Patricia A.
44 REGISTRATION NUMBER: 34,693
45 REFERENCE/DOCKET NUMBER: 2802-D
46 TELECOMMUNICATION INFORMATION:
47 TELEPHONE: 2065870430
48 TELEFAX: 2065870606
49 INFORMATION FOR SEQ ID NO: 4:
50 SEQUENCE CHARACTERISTICS:
51 LENGTH: 519 base pairs
52 TYPE: nucleic acid
53 STRANDEDNESS: single
54 TOPOLOGY: linear
55 MOLECULE TYPE: cDNA
56 HYPOTHETICAL: NO
57 ANTI-SENSE: NO
58 ORIGINAL SOURCE:
59 ORGANISM: HUMAN
60 IMMEDIATE SOURCE:
61 CLONE: CD40 EXTRACELLULAR REGION
62 US-08-770-974-4

Query Match 6.0%; Score 39.2; DB 4; Length 519;
Best Local Similarity 56.1%; Pred. No. 0.06;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 79 TAACACCACTGCTCCCTGCAAGGAGGCACTTCAGATACCTCTCCCGCAGCGC 138
DB 359 TGATACATCTCGAGCGCTGCGCCAGCTGCTTCTCCATGTGTCATCTGCTTGA 418
QY 139 CCGCTGCAGCCCAACACAGGTGTGAGAACCAAGTGTGTGAGGAGAGTCCAGGAC 198
DB 419 AAATGTCACCTTGCAACAGCTGTGAGACCAAGAGACCTGTTGTGCAACAGGAGCGAC 478

FILED DATE: May 24, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/969,703
FILING DATE: October 23, 1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/805,723
FILING DATE: December 5, 1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/783,707
FILING DATE: October 25, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2802-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2065870430
TELEFAX: 2065870606
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: HUMAN
ORGANISM: HUMAN
IMMEDIATE SOURCE:
CLONE: CD40 EXTRACELLULAR REGION
US-08-477-733B-4

Query Match
Best Local Similarity 56.1%; Score 39.2; DB 2; Length 519;
Pred. No. 0.06;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 79 TACACCACTGCGTCCCTGCGACAGGCGCACTTCCAGATCTCTCCCGCCGCGC 138
DB 359 TGATACCACTGCGAGCCCTGCGCCAGTGGCTTCTTCCAGATGTCATCTGCTTCA 418
QY 139 CCGCTCCAGCCCGACAGGAGTGTGAGAACCAAGTGTGTGGAGGCGAGTCCAGGAC 198
DB 419 AAATGTCAACCTGTGAGACAGGTGTGAGACCAAGACCTGTGTGTGCAACAGGCGAC 478
QY 199 TGCCCGAGTCCGA 210
DB 479 AACCAAGACTGA 490

RESULT 4
US-09-088-913A-4
Sequence 4, Application US/09088913A
Patent No. 6087329
GENERAL INFORMATION:
APPLICANT: ARMITAGE, RICHARD
APPLICANT: FANSLAW, WILLIAM
APPLICANT: SPRIGGS, MELANIE
APPLICANT: SRINIVASAN, SUBHASHINI
APPLICANT: GIBSON, MARYLOU
APPLICANT: MORRIS, ARVIA E.
APPLICANT: MCGREW, JEFFERY
TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA

ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/088,913A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/484,624
FILING DATE:
APPLICATION NUMBER: 08/477,733
FILING DATE: June 07, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/969,703
FILING DATE: October 23, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/805,723
FILING DATE: December 5, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/783,707
FILING DATE: October 25, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2802-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2065870430
TELEFAX: 2065870606
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: HUMAN
ORGANISM: HUMAN
IMMEDIATE SOURCE:
CLONE: CD40 EXTRACELLULAR REGION
US-09-088-913A-4

Query Match
Best Local Similarity 56.1%; Score 39.2; DB 3; Length 519;
Pred. No. 0.06;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 79 TACACCACTGCGTCCCTGCGACAGGCGCACTTCCAGATCTCTCCCGCCGCGC 138
DB 359 TGATACCACTGCGAGCCCTGCGCCAGTGGCTTCTTCCAGATGTCATCTGCTTCA 418
QY 139 CCGCTCCAGCCCGACAGGAGTGTGAGAACCAAGTGTGTGGAGGCGAGTCCAGGAC 198
DB 419 AAATGTCAACCTGTGAGACAGGTGTGAGACCAAGACCTGTGTGTGCAACAGGCGAC 478
QY 199 TGCCCGAGTCCGA 210
DB 479 AACCAAGACTGA 490

RESULT 5
US-08-769-819-4
Sequence 4, Application US/08769819
Patent No. 6264951
GENERAL INFORMATION:
APPLICANT: ARMITAGE, RICHARD
APPLICANT: FANSLAW, WILLIAM
APPLICANT: SPRIGGS, MELANIE
APPLICANT: SRINIVASAN, SUBHASHINI
APPLICANT: GIBSON, MARYLOU

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: HUMAN
IMMEDIATE SOURCE:
CLONE: CD40 EXTRACELLULAR REGION
US-08-249-189-4

Query Match
Best Local Similarity 56.1%; Score 39.2; DB 2; Length 519;
Pred. No. 0.06;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 79 TACACACTGCGTCCCTGCAAGGCGACTTCCAGATACCTCTCCCGCCAGCGC 138
DB 359 TGATACCATCTGCGAGCGCTGCCAGTCGCTTCTCCATGTGTCACTGCTTCA 418
QY 139 CCGCTGCCAGCCGACACAGGCTGTGAGAACCAAGTGTGTGAGGCGAGCTCCAGCGC 198
DB 419 AAATGTACCCCTTGCAACAGCTGTGAGACCAAGACCTGTGTGTGCAACAGGCGAGC 478
QY 199 TGCCCACTCCGA 210
DB 479 AACACAGACTGA 490

RESULT 2
US-08-484-624A-4
Sequence 4, Application US/08484624A
Patent No. 5962406
GENERAL INFORMATION:

APPLICANT: ARMSTRONG, RICHARD
APPLICANT: FANSLON, WILLIAM
APPLICANT: SPRIGGS, MELANIE
APPLICANT: SRINIVASAN, SUBHASHINI
APPLICANT: GIBSON, MARYLOU
APPLICANT: MORRIS, ARVIA E.
APPLICANT: MCGREW, JEFFERY
TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,624A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/477,733
FILING DATE: June 07, 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/969,703
FILING DATE: October 23, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/805,723
FILING DATE: December 5, 1991
CLASSIFICATION:
CURRENT APPLICATION DATA:
PRIOR APPLICATION NUMBER: 07/783,707

FILING DATE: October 25, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2802-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2065870430
TELEFAX: 2065870606
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: HUMAN
IMMEDIATE SOURCE:
CLONE: CD40 EXTRACELLULAR REGION
US-08-484-624A-4

Query Match
Best Local Similarity 56.1%; Score 39.2; DB 2; Length 519;
Pred. No. 0.06;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 79 TACACACTGCGTCCCTGCAAGGCGACTTCCAGATACCTCTCCCGCCAGCGC 138
DB 359 TGATACCATCTGCGAGCGCTGCCAGTCGCTTCTCCATGTGTCACTGCTTCA 418
QY 139 CCGCTGCCAGCCGACACAGGCTGTGAGAACCAAGTGTGTGAGGCGAGCTCCAGCGC 198
DB 419 AAATGTACCCCTTGCAACAGCTGTGAGACCAAGACCTGTGTGTGCAACAGGCGAGC 478
QY 199 TGCCCACTCCGA 210
DB 479 AACACAGACTGA 490

RESULT 3
US-08-477-733B-4
Sequence 4, Application US/08477733B
Patent No. 5981724
GENERAL INFORMATION:

APPLICANT: ARMSTRONG, RICHARD
APPLICANT: FANSLON, WILLIAM
APPLICANT: SPRIGGS, MELANIE
APPLICANT: SRINIVASAN, SUBHASHINI
APPLICANT: GIBSON, MARYLOU
APPLICANT: MORRIS, ARVIA E.
APPLICANT: MCGREW, JEFFERY
TITLE OF INVENTION: NOVEL CYTOKINE THAT BINDS CD40
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: MS Word for Apple Power Macintosh, version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,733B
FILING DATE: June 07, 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/249,189

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:11:20 ; Search time 20.9124 Seconds
(without alignments)
9517.464 Million cell updates/sec

Title: US-09-917-372-2_COPY_629_1277

Perfect score: 649

Sequence: 1 cgagtgtaacactgcgagc.....aacctcaccattccc 649

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_NA: *
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/lna/6C.COMB.seq: *
6: /cgn2_6/ptodata/1/lna/6ackfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39.2	6.0	519	2	US-08-249-189-4 Sequence 4, Appli
2	39.2	6.0	519	2	US-08-484-624A-4 Sequence 4, Appli
3	39.2	6.0	519	2	US-08-477-733B-4 Sequence 4, Appli
4	39.2	6.0	519	3	US-09-088-913A-4 Sequence 4, Appli
5	39.2	6.0	519	4	US-08-769-819-4 Sequence 4, Appli
6	39.2	6.0	519	4	US-08-770-974-4 Sequence 4, Appli
7	39.2	6.0	519	4	US-08-770-981-4 Sequence 4, Appli
8	39.2	6.0	519	4	US-09-399-106-4 Sequence 4, Appli
9	39.2	6.0	1004	4	US-09-071-433-85 Sequence 85, Appli
10	39.2	6.0	1004	4	US-09-041-886-26 Sequence 26, Appli
11	39.2	6.0	1004	4	US-08-114-944D-1 Sequence 1, Appli
12	36.4	5.6	459	4	US-09-286-529-7 Sequence 1, Appli
13	36.4	5.6	1164	4	US-08-794-796-1 Sequence 1, Appli
14	36.4	5.6	1347	4	US-09-286-529-18 Sequence 18, Appli
15	36.4	5.6	1859	4	US-09-286-529-19 Sequence 19, Appli
16	35.2	5.4	11282	4	US-09-754-250-3 Sequence 3, Appli
17	35	5.4	382	4	US-08-866-340-29 Sequence 29, Appli
18	35	5.4	382	4	US-09-103-875-35 Sequence 35, Appli
19	34.2	5.3	268	1	US-08-039-137-16 Sequence 16, Appli
20	33.8	5.2	933	6	5340934-12 Patent No. 5340934
21	33	5.1	1057	3	US-09-188-930-18 Sequence 18, Appli
22	33	5.1	1722	4	US-09-142-623-14 Sequence 14, Appli
23	33	5.1	2004	3	US-09-188-930-230 Sequence 230, App
24	32.8	5.1	376	4	US-09-149-476-201 Sequence 201, App
25	32.6	5.0	3282	1	US-08-276-852-154 Sequence 154, App
26	32.6	5.0	3282	1	US-08-276-852-169 Sequence 169, App
27	32.6	5.0	3282	1	US-08-899-575-154 Sequence 154, App

C	28	32.6	5.0	3282	1	US-08-899-575-169	Sequence 169, App
C	29	32.6	5.0	3282	1	US-08-899-575-154	Sequence 154, App
C	30	32.6	5.0	3282	1	US-08-899-575-169	Sequence 169, App
C	31	32.6	5.0	3282	5	PCT-US95-08743-154	Sequence 154, App
C	32	32.6	5.0	3282	5	PCT-US95-08743-169	Sequence 169, App
C	33	32.6	5.0	4723	4	US-09-042-353-370	Sequence 370, App
C	34	32.6	5.0	4723	4	US-08-758-417A-218	Sequence 218, App
C	35	32.6	5.0	4926	4	US-09-042-353-418	Sequence 418, App
C	36	32.6	5.0	4926	4	US-08-758-417A-268	Sequence 268, App
C	37	32.6	5.0	5337	2	US-08-784-512-4	Sequence 4, Appli
C	38	32.6	5.0	5337	4	US-09-176-228-4	Sequence 4, Appli
C	39	32.6	5.0	8614	4	US-09-247-352-5	Sequence 5, Appli
C	40	32.6	5.0	8614	4	US-09-466-635-5	Sequence 5, Appli
C	41	32.6	5.0	13254	1	US-08-276-852-156	Sequence 156, App
C	42	32.6	5.0	13254	1	US-08-276-852-170	Sequence 170, App
C	43	32.6	5.0	13254	1	US-08-899-575-156	Sequence 156, App
C	44	32.6	5.0	13254	1	US-08-899-575-170	Sequence 170, App
C	45	32.6	5.0	13254	1	US-08-899-575-156	Sequence 156, App

ALIGNMENTS

RESULT 1
US-08-249-189-4
Sequence 4, Application US/08249189
Patent No. 5961974
GENERAL INFORMATION:
APPLICANT: ARMITAGE, RICHARD
APPLICANT: FANSLAW, WILLIAM
APPLICANT: SPRIGGS, MELANIE
APPLICANT: SRINIVASAN, SUBHASHINI
APPLICANT: GIBSON, MARYLOU
TITLE OF INVENTION: NOVEL CYTOKINE
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: IMMUNEX CORPORATION
STREET: 51 UNIVERSITY STREET
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Apple Operating System 7.1
SOFTWARE: Microsoft Word for Apple, version 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,189
FILING DATE:
CLASSIFICATION: 514
APPLICATION NUMBER: 07/969,703
FILING DATE: October 23, 1992
CLASSIFICATION: 514
APPLICATION NUMBER: 07/805,723
FILING DATE: December 5, 1991
CLASSIFICATION: 514
APPLICATION NUMBER:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/83,707
FILING DATE: October 25, 1991
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia A.
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2802-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2065870430
TELEFAX: 2065870606
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 base pairs

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAMI0578 row: b column: 10

High quality sequence stop: 673.

FEATURES

source

Location/Qualifiers

1. 700

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4707753"

/clone_lib="NIH_MGC_87"

/tissue_type="mammary adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.383 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 160 a 223 c 189 g 127 t 1 others

ORIGIN

Query Match

57.0%; Score 370; DB 12; Length 700;

Best Local Similarity 99.5%; Pred. No. 2e-82;

Matches 392; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```
QY 257 GGATCGCTGCTCAGAGGCGTCCGAGGAGAGGAGCCCAATCCTGTAGTGAAGCTGG 316
    |||||||
DB 84 GGATCGCTGCTCAGAGGCGTCCGAGGAGAGGAGCCCAATCCTGTAGTGAAGCTGG 316
    |||||||

QY 317 GAGCCTCCGAAGGCCCATCTACTTCCCTGACTTGTGTACAGCCACTGCTACCCATTCT 376
    |||||||
DB 144 GAGCCTCCGAAGGCCCATCTACTTCCCTGACTTGTGTACAGCCACTGCTACCCATTCT 376
    |||||||

QY 377 GGAGATGTTCCCGAGTATCCAGTGGGCTCCCGGAGGCCCACTTTGGAGGAGGGGTG 436
    |||||||
DB 204 GGAGATGTTCCCGAGTATCCAGTGGGCTCCCGGAGGCCCACTTTGGAGGAGGGGTG 436
    |||||||

QY 437 CCGCAACAGCAGAGTCTCTGACCTGACCAAGGAGCCGAGTTGGAACCGGGGAGCAG 496
    |||||||
DB 264 CCGCAACAGCAGAGTCTCTGACCTGACCAAGGAGCCGAGTTGGAACCGGGGAGCAG 496
    |||||||

QY 497 AGCCAGGTGGCCACGCTACCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 556
    |||||||
DB 324 AGCCAGGTGGCCACGCTACCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 556
    |||||||

QY 557 GGCAC-ATCTACATCTACATGAGCAGTACTGGGGGAGCACCAGGCTCTGAGAGCT 615
    |||||||
DB 384 GGCACATCTACATGAGCAGTACTGGGGGAGCACCAGGCTCTGAGAGCT 615
    |||||||

QY 616 CCCAGTACCCCGAGACCTCCATCCCATTTCC 649
    |||||||
DB 444 CCCAGTAC-CCCGAGACCTCCATCCCATTTCC 649
    |||||||
```

Search completed: April 15, 2003, 16:26:25
Job time : 756.63 secs

/note="Organ: ovary; Vector: pOMB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 181 a 297 c 267 g 156 t

ORIGIN

Query Match 58.7%; Score 381; DB 12; Length 901;
Best Local Similarity 99.7%; Pred. No. 3.9e-85;
Matches 392; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 257 GGATCGCTCTCAGAGGCGCTCCGAGGAGAGGAGACCAATCTGTAGCTGAGCTGG 316
|||||
DB 55 GGATCGCTCTCAGAGGCGCT-CCGAGGAGAGGAGACCAATCTGTAGCTGAGAGCTGG 113
|||||
QY 317 GAGCCTCGAAGGCCCATCATCTCCCTGACTGTACAGCACTGCTACCATTTCT 376
|||||
DB 114 GAGCCTCGAAGGCCCATCATCTCCCTGACTGTACAGCACTGCTACCATTTCT 173
|||||
QY 377 GGAGATGTTTCCCATATCACTGAGGCTCCCGAGAGCCCAAGTTTGGAGGAGGGTG 436
|||||
DB 174 GGAGATGTTTCCCATATCACTGAGGCTCCCGAGAGCCCAAGTTTGGAGGAGGGTG 233
|||||
QY 437 CCGCAACAGCAGAGCTCTGTGACCTGACAGGAGGAGCCGAGTTTGAACCCGGGAGCAG 496
|||||
DB 234 CCGCAACAGCAGAGCTCTGTGACCTGACAGGAGGAGCCGAGTTTGAACCCGGGAGCAG 293
|||||
QY 497 AGCCAGGTGGCCAGGCTACCATGTGATGATGATGATGATGATGATGATGATGATGAT 556
|||||
DB 294 AGCCAGGTGGCCAGGCTACCATGTGATGATGATGATGATGATGATGATGATGATGAT 353
|||||
QY 557 GGCACATCTACATCTACATGAGCAGAGTGGGGGAGACACGGGGTCTCTGAGACCTC 616
|||||
DB 354 GGCACATCTACATCTACATGAGCAGAGTGGGGGAGACACGGGGTCTCTGAGACCTC 413
|||||
QY 617 CCAGCTACCCCGAAGCTCATACCCCATTTCCC 649
|||||
DB 414 CCAGCTACCCCGAAGCTCATACCCCATTTCCC 446
|||||

RESULT 14 954 bp mRNA linear EST 20-AUG-2002
B0930450
LOCUS B0930450
DEFINITION AGENCOURT_8946139 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6462494
5', mRNA sequence.
ACCESSION B0930450
VERSION B0930450.1 GI:22345481
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 954)
TITLE NIH-MGC http://mhc.nci.nih.gov/
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2652 row: h column: 15
High quality sequence stop: 576.
Location/Qualifiers
1..954

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6462494"
/clone_11b="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOMB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library."

BASE COUNT 205 a 320 c 278 g 151 t

ORIGIN

Query Match 58.6%; Score 380.4; DB 14; Length 954;
Best Local Similarity 99.7%; Pred. No. 5.6e-85;
Matches 381; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 268 CAAGAGCGCTCCGAGGAGAGGAGACCAATCTGTAGCTGAGAGCTGGAGCTCCGAA 327
|||||
DB 1 CAAGAGCGCTCCGAGGAGAGGAGACCAATCTGTAGCTGAGAGCTGGAGCTCCGAA 60
|||||
QY 328 GGGCCATCATCTTCCCTGACTGTGATGATGATGATGATGATGATGATGATGATGAT 387
|||||
DB 61 GGGCCATCATCTTCCCTGACTGTGATGATGATGATGATGATGATGATGATGATGAT 120
|||||
QY 388 CCCAGTATCATCTGAGGCTCCCGAGAGCCCAAGTTTGAAGGAGGGGTCCGACAGCA 447
|||||
DB 121 CCAGTATCATCTGAGGCTCCCGAGAGCCCAAGTTTGAAGGAGGGGTCCGACAGCA 180
|||||
QY 448 GAGTCTCTGAGAGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 507
|||||
DB 181 GAGTCTCTGAGAGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
|||||
QY 508 CCAGGTTACCAATGAGCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 567
|||||
DB 241 CCAGGTTACCAATGAGCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 300
|||||
QY 568 CATCTCAATGAGCAGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 627
|||||
DB 301 CATCTCAATGAGCAGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
|||||
QY 628 CGAAGCTCATACCCCATTTCCC 649
|||||
DB 361 CGAAGCTCATACCCCATTTCCC 382
|||||

RESULT 15 700 bp mRNA linear EST 10-APR-2001
BG576875
LOCUS BG576875
DEFINITION 602599233F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4707753 5',
mRNA sequence.
ACCESSION BG576875
VERSION BG576875.1 GI:13584528
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 700)
TITLE NIH-MGC http://mhc.nci.nih.gov/
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

Best Local Similarity 100.0%; Pred. No. 3.4e-85;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 257 GATCGCTGCTCAAGAGCGCTCCGACAGAGAGAGACCAATCTCTAGCTGGAAGCTGG 316
      |||||||
Db 260 GGATTCCTGCTCAAGAGCGCTCCGACAGAGAGAGACCAATCTCTAGCTGGAAGCTGG 319
      |||||||
QY 317 GAGCCTCCGAAGGCCATTCATCTCTGACTTGTAGACGACCATGCTACCATTTCT 376
      |||||||
Db 320 GAGCCTCCGAAGGCCATTCATCTCTGACTTGTAGACGACCATGCTACCATTTCT 379
      |||||||
QY 377 GGAGATTTTCCCGAGTATCCACTGGGCTCCCGACAGCCCGCATTTTGGAGGCGGGGTG 436
      |||||||
Db 380 GGAGATTTTCCCGAGTATCCACTGGGCTCCCGACAGCCCGCATTTTGGAGGCGGGGTG 439
      |||||||
QY 437 CCGCAACAGAGAGATCTCTGAGACTGACGAGGAGGAGGAGTGGAAACCGGGGAGCAG 496
      |||||||
Db 440 CCGCAACAGAGAGATCTCTGAGACTGACGAGGAGGAGGAGTGGAAACCGGGGAGCAG 499
      |||||||
QY 497 AGCCAGGTGGCCACGGTACCAATGGCATTCATGTACCGCGGGGTCTATGACTATCACT 556
      |||||||
Db 500 AGCCAGGTGGCCACGGTACCAATGGCATTCATGTACCGCGGGGTCTATGACTATCACT 559
      |||||||
QY 557 GGCACATCTACATCTACATGACAGTACTGGGGGAGCACCGGGTCTCTGGAGACCTC 616
      |||||||
Db 560 GGCACATCTACATCTACATGACAGTACTGGGGGAGCACCGGGTCTCTGGAGACCTC 619
      |||||||
QY 617 CCAGTACCCCCGAACTCCA 637
      |||||||
Db 620 CCAGTACCCCCGAACTCCA 640
      |||||||

```

RESULT 12
B1829057 714 bp mRNA linear EST 04-OCT-2001
LOCUS 603079147F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5171122 5',
DEFINITION mRNA sequence.
ACCESSION B1829057
VERSION B1829057.1 GI:15940607
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM11425 row: m column: 11
High quality sequence stop: 711.
Location/Qualifiers

FEATURES

1. 714
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5171122"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
BASE COUNT 152 a 231 c 202 g 129 t

Query Match 58.7%; Score 381; DB 13; Length 714;
Best Local Similarity 99.7%; Pred. No. 3.5e-85;
Matches 392; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

QY 257 GATCGCTGCTCAAGAGCGCTCCGACAGAGAGAGACCAATCTCTAGCTGGAAGCTGG 316
      |||||||
Db 81 GGATTCCTGCTCAAGAGCGCTCCGACAGAGAGAGACCAATCTCTAGCTGGAAGCTGG 140
      |||||||
QY 317 GAGCCTCCGAAGGCCATTCATCTCTGACTTGTAGACGACCATGCTACCATTTCT 376
      |||||||
Db 141 GAGCCTCCGAAGGCCATTCATCTCTGACTTGTAGACGACCATGCTACCATTTCT 200
      |||||||
QY 377 GGAGATTTTCCCGAGTATCCACTGGGCTCCCGACAGCCCGCATTTTGGAGGCGGGGTG 436
      |||||||
Db 201 GGAGATTTTCCCGAGTATCCACTGGGCTCCCGACAGCCCGCATTTTGGAGGCGGGGTG 260
      |||||||
QY 437 CCGCAACAGAGAGATCTCTGAGACTGACGAGGAGGAGGAGTGGAAACCGGGGAGCAG 496
      |||||||
Db 261 CCGCAACAGAGAGATCTCTGAGACTGACGAGGAGGAGGAGTGGAAACCGGGGAGCAG 319
      |||||||
QY 497 AGCCAGGTGGCCACGGTACCAATGGCATTCATGTACCGCGGGGTCTATGACTATCACT 556
      |||||||
Db 320 AGCCAGGTGGCCACGGTACCAATGGCATTCATGTACCGCGGGGTCTATGACTATCACT 379
      |||||||
QY 557 GGCACATCTACATCTACATGACAGTACTGGGGGAGCACCGGGTCTCTGGAGACCTC 616
      |||||||
Db 380 GGCACATCTACATCTACATGACAGTACTGGGGGAGCACCGGGTCTCTGGAGACCTC 439
      |||||||
QY 617 CCAGTACCCCCGAACTCCA 649
      |||||||
Db 440 CCAGTACCCCCGAACTCCA 472
      |||||||

```

RESULT 13
BE740143 901 bp mRNA linear EST 15-SEP-2000
LOCUS 601595085F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949148 5',
DEFINITION mRNA sequence.
ACCESSION BE740143
VERSION BE740143.1 GI:10154135
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: LLM813 row: i column: 21
High quality sequence stop: 745.
Location/Qualifiers

FEATURES

1. 901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949148"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"

Query Match	59.8%	Score 388	DB 10	Length 427
Best Local Similarity	100.0%	Pred. No. 5e-87		
Matches 388	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	262	GCTCTCTCAAGAGCGCTCCGACAGAGAGAGACCCATCTGTAGCTGGAAAGCTGGAGCC	321	
Db	1	GCTGCTCAAGAGCGCTCCGACAGAGAGAGACCCATCTGTAGCTGGAAAGCTGGAGCC	60	

Query Match	58.7%	Score 381;	DB 12;	Length 640;
-------------	-------	------------	--------	-------------

```

Db 432 GGAACATCTACATCTACATCTGAGCCAGTACTGGGGGAGACACCGGCTCTGAGACTC 491
Qy 617 CCAGCTACCCCCGAGACCTCCATACCCCATTC 649
Db 492 CCAGCTACCCCCGAGACCTCCATACCCCATTC 524

RESULT 8
LOCUS BG036632 920 bp mRNA linear EST 24-JAN-2001
DEFINITION 602326834F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4428225 5',
mRNA sequence.
ACCESSION BG036632
VERSION BG036632.1 GI:12432013
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 920)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM0178 row: c column: 10
High quality sequence stop: 695.
Location/Qualifiers
1. 920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4428225"
/clone_11b="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: prostate; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 187 a 304 c 270 g 159 t
ORIGIN
Query Match 60.6%; Score 393; DB 12; Length 920;
Best Local Similarity 100.0%; Pred. No. 3.8e-88;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGAGACCAATCTGTAGTGAAGCTGG 316
Db 52 GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGAGACCAATCTGTAGTGAAGCTGG 111
Qy 317 GAGCTCCGGAAGGCCCATCTACTCTGCTGACTTGGTGAAGCCATGCTCAATTTCT 376
Db 112 GAGCTCCGGAAGGCCCATCTACTCTGCTGACTTGGTGAAGCCATGCTCAATTTCT 171
Qy 377 GGAATGTTTCCAGATCCAGTACAGTGGGCTCCGAGGCCAGCCAGTTTGGAGGAGGGTG 436
Db 172 GGAATGTTTCCAGATCCAGTACAGTGGGCTCCGAGGCCAGCCAGTTTGGAGGAGGGTG 231
Qy 437 CCGCAACAGCAGAGTCTCTGACCTGACACGAGAGCCAGTGTGAACCCGGGGAGCAG 496
Db 232 CCGCAACAGCAGAGTCTCTGACCTGACACGAGAGCCAGTGTGAACCCGGGGAGCAG 291
Qy 497 AGCCAGTGGCCACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 556

```

```

Db 292 AGCCAGTGGCCCGCCAGCTACCATGATGCATGACCGGGGGTCTATGACTACAT 351
Qy 557 GGAACATCTACATCTACATCTGAGCCAGTACTGGGGGAGACACCGGCTCTGAGACTC 616
Db 352 GGAACATCTACATCTACATCTGAGCCAGTACTGGGGGAGACACCGGCTCTGAGACTC 411
Qy 617 CCAGCTACCCCCGAGACCTCCATACCCCATTC 649
Db 412 CCAGCTACCCCCGAGACCTCCATACCCCATTC 444

RESULT 9
LOCUS B1821009 954 bp mRNA linear EST 04-OCT-2001
DEFINITION 603035664F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176935 5',
mRNA sequence.
ACCESSION B1821009
VERSION B1821009.1 GI:15932559
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1440 row: o column: 16
High quality sequence stop: 758.
Location/Qualifiers
1. 954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5176935"
/clone_11b="NIH_MGC_115"
/lab_host="DH10B"
/notes="Organ: pooled brain, lung, testis; Vector:
pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT 212 a 313 c 281 g 148 t
ORIGIN
Query Match 60.6%; Score 393; DB 13; Length 954;
Best Local Similarity 100.0%; Pred. No. 3.9e-88;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGAGACCAATCTGTAGTGAAGCTGG 316
Db 14 GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGAGACCAATCTGTAGTGAAGCTGG 73
Qy 317 GAGCTCCGGAAGGCCCATCTACTCTGCTGACTTGGTGAAGCCATGCTCAATTTCT 376
Db 74 GAGCTCCGGAAGGCCCATCTACTCTGCTGACTTGGTGAAGCCATGCTCAATTTCT 133
Qy 377 GGAATGTTTCCAGATCCAGTACAGTGGGCTCCGAGGCCAGCCAGTTTGGAGGAGGGTG 436
Db 134 GGAATGTTTCCAGATCCAGTACAGTGGGCTCCGAGGCCAGCCAGTTTGGAGGAGGGTG 193

```

DEFINITION 601122062F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3346311 5', mRNA sequence.

ACCESSION BE275240

VERSION BE275240.1 GI:9150193

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 868)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L10CM135 row: K column: 16
High quality sequence stop: 746.
Location/Qualifiers

1. 868

source /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3346311"
/clone_lib="NIH_MGC_20"
/tissue_type="melanocytic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site.1: XhoI; Site.2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene, Berkeley) and Superscript II RT (Life Technologies)."

BASE COUNT 170 a 286 c 253 g 159 t

ORIGIN

Query Match 60.6%; Score 393; DB 10; Length 868;
Best Local Similarity 100.0%; Pred. No. 3.7e-88;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

257 GGATCGCTGCTCAAGAGGCGTCCGAGAGAGAGACCAATCTGTAGCTGGAAGCTGG 316
|||||
138 GGATGCTGCTCAAGAGGCGTCCGAGAGAGAGACCAATCTGTAGCTGGAAGCTGG 197
|||||
317 GAGCTTCGGAAGGCCCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCATTTCT 376
|||||
198 GAGCTTCGGAAGGCCCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCATTTCT 257
|||||
377 GAGATGTTTTCCCAATATCCACTGGGCTCCCGACAGCCCGAGTTTGGAGGCGAGGGTG 436
|||||
258 GAGATGTTTTCCCAATATCCACTGGGCTCCCGACAGCCCGAGTTTGGAGGCGAGGGTG 317
|||||
437 CCGCAACAGCAGAGTCTCTGAGCTGACACAGGAGCCGACATTTGGAACCCGGGAGCAG 496
|||||
318 CCGCAACAGCAGAGTCTCTGAGCTGACACAGGAGCCGACATTTGGAACCCGGGAGCAG 377
|||||
497 AGCCAGGTGGCCCGACGGTACCAATGGCATTTGTCACCGGGGGTCTATGACTATCACT 556
|||||
378 AGCCAGGTGGCCCGACGGTACCAATGGCATTTGTCACCGGGGGTCTATGACTATCACT 437
|||||
557 GGCACATCTACATCTACAATGGACAGTACGAGGGGGAGCCACCGGGTCTGAGAACCTC 616
|||||
438 GGCACATCTACATCTACAATGGACAGTACGAGGGGGAGCCACCGGGTCTGAGAACCTC 497
|||||
617 CCAGCTACCCCGACGATCTCATATCCCATTTCCC 649
|||||
498 CCAGCTACCCCGACGATCTCATATCCCATTTCCC 530

RESULT 7

LOCUS B0957322 895 bp mRNA linear EST 21-AUG-2002

DEFINITION AGENCOURT_8777754 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384424 5', mRNA sequence.

ACCESSION B0957322

VERSION B0957322.1 GI:22372800

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 895)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L10CM2579 row: g column: 17
High quality sequence stop: 645.
Location/Qualifiers

1. 895

source /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6384424"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pORF7; Site.1: XhoI; Site.2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 190 a 291 c 252 g 159 t 3 others

ORIGIN

Query Match 60.6%; Score 393; DB 14; Length 895;
Best Local Similarity 100.0%; Pred. No. 3.8e-88;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

257 GGATCGCTGCTCAAGAGGCGTCCGAGAGAGAGACCAATCTGTAGCTGGAAGCTGG 316
|||||
132 GGATGCTGCTCAAGAGGCGTCCGAGAGAGAGACCAATCTGTAGCTGGAAGCTGG 191
|||||
317 GAGCTTCGGAAGGCCCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCATTTCT 376
|||||
192 GAGCTTCGGAAGGCCCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCATTTCT 251
|||||
377 GAGATGTTTTCCCAATATCCACTGGGCTCCCGACAGCCCGAGTTTGGAGGCGAGGGTG 436
|||||
252 GAGATGTTTTCCCAATATCCACTGGGCTCCCGACAGCCCGAGTTTGGAGGCGAGGGTG 311
|||||
437 CCGCAACAGCAGAGTCTCTGAGCTGACACAGGAGCCGACATTTGGAACCCGGGAGCAG 496
|||||
312 CCGCAACAGCAGAGTCTCTGAGCTGACACAGGAGCCGACATTTGGAACCCGGGAGCAG 371
|||||
497 AGCCAGGTGGCCCGACGGTACCAATGGCATTTGTCACCGGGGGTCTATGACTATCACT 556
|||||
372 AGCCAGGTGGCCCGACGGTACCAATGGCATTTGTCACCGGGGGTCTATGACTATCACT 431
|||||
557 GGCACATCTACATCTACAATGGACAGTACGAGGGGGAGCCACCGGGTCTGAGAACCTC 616

OY 292 ACCCAATCTGTAGCTGGAAGCTGGAGCCTCCGAAAGGCCATTCATCTTCCCTGACT 351
 DB 301 ACCCAATCTGTAGCTGGAAGCTGGAGCCTCCGAAAGGCCATTCATCTTCCCTGACT 360
 OY 352 GGTACAGCCACTGCTACCTATTTCTGGAGATGTTTCCAGATCCACTGGGCTCCCGC 411
 DB 361 GGTACAGCCACTGCTACCTATTTCTGGAGATGTTTCCAGATCCACTGGGCTCCCGC 420
 OY 412 AGCCCAAGTTTGGAGGAGGCTGGCCGCAACAGCAGATCTCTGGACCTGACAGGGA 471
 DB 421 AGCCCAAGTTTGGAGGAGGCTGGCCGCAACAGCAGATCTCTGGACCTGACAGGGA 480
 OY 472 GCGCAGTTTGAACCCGCGGAGACAGACAGGTCGCCACAGGTACCAATGGCATTCATGT 531
 DB 481 GCGCAGTTTGAACCCGCGGAGACAGACAGGTCGCCACAGGTACCAATGGCATTCATGT 540
 OY 532 CACCGCGGGTCTATGATCATCTGCAACATCTACATCTACATGAGCAGTACTGG 591
 DB 541 CACCGCGGGTCTATGATCATCTGCAACATCTACATCTACATGAGCAGTACTGG 600
 OY 592 GGGACACCGGGTCTGAGAGACCTCCAGCTACCCCGAAGCTCATACCCCATTCGCC 649
 DB 601 GGGACACCGGGTCTGAGAGACCTCCAGCTACCCCGAAGCTCATACCCCATTCGCC 658
 RESULT 3
 BI769498 900 bp mRNA linear EST 25-SEP-2001
 LOCUS 603059047F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5208394 5'
 DEFINITION mRNA sequence.
 ACCESSION BI769498
 VERSION BI769498.1 GI:15761076
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 900)
 NIH-MGC http://mgc.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLM11522 row: n column: 11
 High quality sequence stop: 854.
 Location/Qualifiers
 1. 900
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5208394"
 /clone_id="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
 Site:1: NotI; Site:2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH-MGC Library."
 BASE COUNT 199 a 298 c 246 g 157 t
 ORIGIN
 Query Match 64.9%; Score 421.4; DB 13; Length 900;

Best Local Similarity 83.4%; Pred. No. 2 8e-95;
 Matches 551; Conservative 0; Mismatches 1; Indels 109; Gaps 2;
 OY 97 CTGCAAGCCAGGACCTTCAGAAATACCTCTCCCGCAGCCCGCTGCCAGCCACAC 156
 DB 1 CTGCAAGCCAGGACCTTCAGAAATACCTCTCCCGCAGCCCGCTGCCAGCCACAC 60
 OY 157 CAGGTGTAGAAACCAAGTCTGTGGAGGAGCTCCAGGACCTGCCAGTCCGACACAC 216
 DB 61 CAGGTGTAGAAACCAAGTCTGTGGAGGAGCTCCAGGACCTGCCAGTCCGACACAC 120
 OY 217 CTGCAAAATCCATTAGAGCCACTGCCCCAGAGATGTCA----- 256
 DB 121 CTGCAAAATCCATTAGAGCCACTGCCCCAGAGATGTCAAGAACATGTGATCTGCC 180
 OY 257 ----- 256
 DB 181 CGTCTGTGCGCAGCTGCGCTTCTTGTGCTCTGCTCCACCGTCTTCTGCTGATCTGAA 240
 OY 257 -----GGATCGCTGCTCAAGAGCGCTCCGACAGGAGA 288
 DB 241 GAGCAACCTTCTCTGTGAGGAAACTGGATGCTGCTCAAGAGCGCTCCGACAGGAGA 300
 OY 289 GAGCAACCTTCTCTGTGAGGAAACTGGATGCTGCTCAAGAGCGCTCCGACAGGAGA 348
 DB 301 GGGACCCATCTGTAGCTGGAAGCTGGAGGCTCCGAAAGGCCATTCATCTTCCCTGA 360
 OY 349 CTTGTGTAGCCACTGCTACCTATTTCTGGAGATGTTTCCAGATTCACACTGGGCTCC 408
 DB 361 CTTGTGTAGCCACTGCTACCTATTTCTGGAGATGTTTCCAGATTCACACTGGGCTCC 420
 OY 409 CCGAGCCCAAGTTTGGAGGAGGCTGCCGCAAGCAGAGATCTCTGGACCTGACAG 468
 DB 421 CCGAGCCCAAGTTTGGAGGAGGCTGCCGCAAGCAGAGATCTCTGGACCTGACAGCA- 479
 OY 469 GGAAGCCGAGTTTGAACCCGCGGAGACAGGAGGCTGCCAGGTTCCCAATGGGATTA 528
 DB 480 GGAAGCCGAGTTTGAACCCGCGGAGACAGGAGGCTGCCAGGTTCCCAATGGGATTA 539
 OY 529 TGTACACCGCGGCTATGATCATCTGCAACATCTCATCTCAATGAGCAGTACT 588
 DB 540 TGTACACCGCGGCTATGATCATCTGCAACATCTCATCTCAATGAGCAGTACT 599
 OY 589 GGGGGAGCACCGGGTCTGAGAGACCTCCAGCTACCCCGAAGCTCATACCCCATTC 648
 DB 600 GGGGGAGCACCGGGTCTGAGAGACCTCCAGCTACCCCGAAGCTCATACCCCATTC 659
 OY 649 C 649
 DB 660 C 660
 RESULT 4
 BM764302 538 bp mRNA linear EST 04-MAR-2002
 LOCUS K-EST0045797 S125NU216 Homo sapiens cDNA clone S125NU216-16-E09 5'
 DEFINITION mRNA sequence.
 ACCESSION BM764302
 VERSION BM764302.1 GI:19093917
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 538)
 Kim, N.S., Hahn, Y., Oh, J. H., Lee, J. Y., Ahn, H. Y., Chu, M. Y., Kim, M. R.,
 Oh, K. J., Cheong, J. E., Sohn, H. Y., Kim, J. M., Park, H. S., Kim, S. and
 Kim, Y. S.
 21C Frontier Korean EST Project 2001
 TITLE Unpublished (2002)
 JOURNAL Contact: Kim YS
 COMMENT Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:518648"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/Note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb.
Insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC Library."

BASE COUNT      156 a      233 c      198 g      141 t
ORIGIN
Query Match
Best Local Similarity 84.3%; Pred. No. 1.3e-105;
Matches 592; Conservative 0; Mismatches 1; Indels 109; Gaps 2;

QY 1 CGAGGTACACACTGCGAGCTTCTTCTGACTGCGCGCTGCGACTGAAGCCGAGCTCAA 60
    |||||||
Db 28 CGAGGTACACACTGCGAGCTTCTTCTGACTGCGCGCTGCGACTGAAGCCGAGCTCAA 87
    |||||||

QY 61 AGATGAAGTTGGGAGGTTAAACAACACGCGTCCCTGGAAGGAGGAGCTTCCAGAA 120
    |||||||
Db 88 AGATGAAGTTGGGAGGTTAAACAACACGCGTCCCTGGAAGGAGGAGCTTCCAGAA 147
    |||||||

QY 121 TACCTCTCCCGAGGCGCGCTGCCAGCCCAACAGGTGTGAGAACCAAGGTCGTGT 180
    |||||||
Db 148 TACCTCTCCCGAGGCGCGCTGCCAGCCCAACAGGTGTGAGAACCAAGGTCGTGT 207
    |||||||

QY 181 GGAGGACGCTCCAGGACCTGCCAGACACAACTGCAAAAATTCATTAGAGCCACT 240
    |||||||
Db 208 GGAGGACGCTCCAGGACCTGCCAGACACAACTGCAAAAATTCATTAGAGCCACT 267
    |||||||

QY 241 GCGCCCGAGATGTCA----- 256
    |||||||
Db 268 GCGCCCGAGATGTCAAGAACCAATGCTGATGTGGCGCTTCTGCTGCCACTGCGCTTCT 327
    |||||||

QY 257 ----- 256
    |||||||
Db 328 TCTGCTCCTTGGCAGCGTCTTCTGCTGATGTGAAGAGCCACCTTCTCTGACAGAA 387
    |||||||

QY 257 -----GATGCGTCTCTCAAGAGCGCTCCGACAGGAGGAGCAATCTGTAGCTGGAAG 312
    |||||||
Db 388 ACTGGGATCGCTCTCTCAAGAGCGCTCCGACAGGAGGAGCAATCTGTAGCTGGAAG 447
    |||||||

QY 313 CTGGGAGCTCCGAAAGGCCATCATCTCCGACTGTGTCAGGCAATGCTTACCCAT 372
    |||||||
Db 448 CTGGGAGCTCCGAAAGGCCATCATCTCCGACTGTGTCAGGCAATGCTTACCCAT 507
    |||||||

QY 373 TTCTGAGATGTTCCCACTATCCATGAGGTCTCCCGCAGCCCAAGTTTGGAGGAGG 432
    |||||||
Db 508 TTCTGAGATGTTCCCACTATCCATGAGGTCTCCCGCAGCCCAAGTTTGGAGGAGG 567
    |||||||

QY 433 GTGGCGGACACAGAGTCTCTGAGACTGACGAGGAGCCGAGTTGGAAACCCGGGGA 492
    |||||||
Db 568 GTGGCGGACACAGAGTCTCTGAGACTGACGAGGAGCCGAGTTGGAAACCCGGGGA 626
    |||||||

QY 493 GCAGAGCAGGTGGCCAGGTCACCAATGATTCATGTACCGGGGGGCTGTGACTAT 552
    |||||||
Db 627 GCAGAGCAGGTGGCCAGGTCACCAATGATTCATGTACCGGGGGGCTGTGACTAT 686
    |||||||

QY 553 CACTGCACATCTACATCTACAAATGAGACAGTACTGGGGGG 594
    |||||||
Db 687 CACTGCACATCTACATCTACAAATGAGACAGTACTGGGGGG 728
    |||||||

RESULT 2

```

```

B0898015      968 bp      mRNA      linear      EST 16-AUG-2002
LOCUS
DEFINITION
AGENCOURT_8125942 Lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6177574 5', mRNA sequence.
ACCESSION
B0898015
VERSION
B0898015.1 GI:22290029
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 968)
AUTHORS
NIH-MGC http://mhc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapds@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1355 row: 1 column: 23
High quality sequence stop: 615.
Location/Qualifiers
1. 968
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6177574"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/Note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGCTCCG-3' and
5'-GACTACTTCTAGATCCGCGAGCGCCGCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

BASE COUNT      219 a      320 c      260 g      168 t      1 others
ORIGIN
Query Match
Best Local Similarity 83.4%; Pred. No. 1.6e-97;
Matches 549; Conservative 0; Mismatches 1; Indels 108; Gaps 1;

QY 100 CAAGGACGGGCACTCCAGATCTCTCCCGAGCGCGCTGCCAGCCCAACAG 159
    |||||||
Db 1 CAAGGACGGGCACTCCAGATCTCTCCCGAGCGCGCTGCCAGCCCAACAG 60
    |||||||

QY 160 GTGTGAACAACAGTGTGTGTGAGGAGCTCCAGGACGCTCCGAGCCGACACAACTG 219
    |||||||
Db 61 GTGTGAACAACAGTGTGTGTGAGGAGCTCCAGGACGCTCCGAGCCGACACAACTG 120
    |||||||

QY 220 CAAATAATCCATTAGAGCACTGCCCCAGAGATGTCA----- 256
    |||||||
Db 121 CAAATAATCCATTAGAGCACTGCCCCAGAGATGTCAAGAACCATGTGATGTGGCCGT 180
    |||||||

QY 257 ----- 256
    |||||||
Db 181 TCTGTGCGACATGCGCTTCTTCTGCTCTTCCACCGCTTCTCTGCAATCTGGAAGAG 240
    |||||||

QY 257 -----GGATGCTGTCTCAAGAGCGCTGCCAGGAGGAGG 291
    |||||||
Db 241 CCACCTTCTCTGCGAGGAACACTGGGATGCTGCTCAAGAGGCGTCCGAGGAGAGGG 300
    |||||||

```

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:09:45 ; Search time 749.63 Seconds
(without alignments)
14021.414 Million cell updates/sec

Title: US-09-917-372-2_COPY_629_1277

Perfect score: 649

Sequence: 1 cgaagtacacactgcagac.....aacctccatcccatcc 649

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 segs, 8097743376 residues 32308132

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estum:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	462.4	71.2	728	13	BT65703 603046670
2	430.4	66.3	968	14	B0898015 AGENCOURT
3	421.4	64.9	900	13	BT69498 603059047
4	418	64.4	538	14	BM764302 K-EST0045
5	393	60.6	814	9	AU118203
6	393	60.6	868	10	BE275240 601122062

7	393	60.6	895	14	B0957322
8	393	60.6	920	12	B036632
9	393	60.6	954	13	B1821009
10	388	59.8	427	10	AM630662
11	381	58.7	640	12	BG491235
12	381	58.7	714	13	B1829057
13	381	58.7	901	12	B1829057
14	380.4	58.6	954	14	B0930450
15	370	57.0	700	12	B0376875
16	355	54.7	749	12	BE789430
17	347	53.5	772	12	BE765684
18	346.6	53.0	579	10	BE207591
19	344	53.0	735	12	BG331666
20	343.8	53.0	606	13	B1905922
21	342	52.7	598	12	BG403820
22	322	49.6	801	12	BF792867
23	316	48.7	458	14	BM766807
24	316	48.7	501	14	BM766833
25	316	48.7	537	14	BM766352
26	316	48.7	572	14	BM767287
27	316	48.7	666	14	BM767318
28	314.4	48.4	634	14	BM746911
29	311	47.9	598	14	BM723050
30	301.8	46.5	360	14	BM766382
31	293.4	45.2	386	9	AA099514
32	282	43.5	598	14	B0305454
33	275	42.4	506	14	BM855039
34	274.4	42.3	276	14	BM784961
35	274	42.2	808	13	B1818411
36	273.4	42.1	466	14	B0082995
37	269.8	41.6	426	14	BM856797
38	263.4	40.6	416	14	B0083004
39	263	40.5	479	14	BM741617
40	263	40.5	541	14	BM846805
41	263	40.5	560	14	BM847170
42	263	40.5	594	14	BM763326
43	263	40.5	841	12	BM7650836
44	263	40.5	968	14	B0653019
45	263	40.5	993	9	AL513838

ALIGNMENTS

RESULT 1
LOCUS BT65703 728 bp mRNA linear EST 25-SEP-2001
DEFINITION 603046670P1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186848 5',
mRNA sequence.
ACCESSION BT65703
VERSION BT65703.1 GI:15757281
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 728)
AUTHORS NIH-MGC
TITLE Unpublished (1999)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Stransberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMN at:
http://image.llnl.gov
Plate: L1AM11466 row: 1 column: 17
High quality sequence stop: 718.
Location/Qualifiers
1..728

FEATURES
source

Db 524 AAATGTCACCTGTGACAACTGTGAGACCAAGACCTGTTGTGCAACAGGCGAGGCAC 583
QY 199 TGGCCAGTCGA 210
Db 584 AAACAAGACTGA 595

Search completed: April 15, 2003, 15:16:24
Job time : 131.963 secs

TOLESTH
MAY 10 2003
12:00 PM
12:00 PM

Best Local Similarity 56.1%; Pred. No. 1.3;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 79 TACACACCACTGCTCCCTGCAAGGCGACACTTCCAGATACCTCTCCCGCAGGC 138
DB 464 TGATACCATCTGCGAGCCCTGCCAGTCTTCTCCAGTGTATCTGCTTTGCA 523
QY 139 CCGCTGCCAGCCCGACACAGGTGTGAGAACCAAGTGTGTGAGGCGACTCCAGGCAC 198
DB 524 AAAATGTCACTTGAGACCAAGCTGTGAGACCAAGACCTGTGTGTGCAACAGGCGAC 583
QY 199 TGCCAGTCCGA 210
DB 584 AACACAGACTGA 595

RESULT 14
AAQ47341
ID AAQ47341 standard; cDNA; 1004 BP.
XX
AC AAQ47341;
XX
DT 07-FEB-1994 (first entry)
XX
DE CD40 coding sequence.
XX
KM Receptor; ligand; B-cell; T-cell; allergy; autoimmunity; antibody;
XX ss.
XX OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 48..876
FT CDS /*lag= a
FT /product= CD40CR Receptor.
XX
PN EP555880-A.
XX
PD 18-AUG-1993.
XX
PE 12-FEB-1993; 93EP-0102279.
XX
PR 14-FEB-1992; 92US-0835799.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
PA (DART-) DARTMOUTH COLLEGE.
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Aruffo AA, Ledbetter JA, Noell R, Stamenkovic I;
PI Noelle R;
XX
XX WPI; 1993-260142/33.
DR P-PSDB; AAR38859.
XX
PT CD40CR receptor and its' ligands - used to inhibit B-cell
XX activation in allergy and auto-immune disease
XX
XX Disclosure: Figure 8a; 21pp; English.
XX
XX The CD40CR receptor is a counter receptor for the CD40 B-cell
XX antigen. It is also a receptor for ligands (sometimes fusion
XX molecules) comprising part of the CD40 protein. A soluble
XX CD40/immunoglobulin fusion protein is able to inhibit helper T-cell
XX mediated B-cell activation by binding to the CD40 receptor on
XX T-cell membranes. Purified receptor provides a means of
XX controlling B-cell activation which may be useful in the treatment
XX of allergy and autoimmune disease.
SQ Sequence 1004 BP; 230 A; 314 C; 259 G; 201 T; 0 other;

Query Match 6.0%; Score 39.2; DB 14; Length 1004;
Best Local Similarity 56.1%; Pred. No. 1.3;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 79 TACACACCACTGCTCCCTGCAAGGCGACACTTCCAGATACCTCTCCCGCAGGC 138
DB 464 TGATACCATCTGCGAGCCCTGCCAGTCTTCTCCAGTGTATCTGCTTTGCA 523
QY 139 CCGCTGCCAGCCCGACACAGGTGTGAGAACCAAGTGTGTGAGGCGACTCCAGGCAC 198
DB 524 AAAATGTCACTTGAGACCAAGCTGTGAGACCAAGACCTGTGTGTGCAACAGGCGAC 583
QY 199 TGCCAGTCCGA 210
DB 584 AACACAGACTGA 595

RESULT 15
AAT14706
ID AAT14706 standard; cDNA; 1004 BP.
XX
AC AAT14706;
XX
DT 30-OCT-1996 (first entry).
XX
DE Human CD40 antigen cDNA.
XX
KM Cell surface antigen; cloning; immunoselection; immunotherapy;
XX therapy; diagnosis; vector; CD40; COS; lymphocyte; ss.
XX
OS Homo sapiens.
XX
PN US5506126-A.
XX
PD 09-APR-1996.
XX
PE 25-FEB-1988; 88US-0160416.
XX
PR 01-DEC-1992; 92US-0983647.
XX
PR 25-FEB-1988; 88US-0160416.
XX
PR 13-JUL-1989; 89US-0379076.
XX
PR 13-JUL-1990; 90US-0553759.
XX
PR 18-OCT-1993; 93US-0139273.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Aruffo A, Seed B;
XX
PI WPI; 1996-200279/20.
XX
PT Cloning of cDNA encoding cell surface antigen - useful for isolation
XX of diagnostic and therapeutic proteins
XX
XX Example 8; Fig 16; 79pp; English.
XX
XX A cDNA clone (AAT14714) codes for human cell surface antigen CD40.
XX It was isolated using a rapid immunoselection method in which
XX a cDNA library is constructed in mammalian (e.g. COS) cells using
XX novel expression vectors (see also AAT14702 and AAT14705), and cells
XX expressing the antigen are selected using antibody-coated plates
XX (panning). This immunoselection cloning method, developed to clone
XX genes for cell surface antigens of human lymphocytes (see also
XX AAT14703-04 and AAT14706-26), has general appln. Cell surface antigens
XX are obtd. for diagnostic and therapeutic use.
SQ Sequence 1004 BP; 230 A; 297 C; 276 G; 201 T; 0 other;

Query Match 6.0%; Score 39.2; DB 17; Length 1004;
Best Local Similarity 56.1%; Pred. No. 1.3;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 79 TACACACCACTGCTCCCTGCAAGGCGACACTTCCAGATACCTCTCCCGCAGGC 138
DB 464 TGATACCATCTGCGAGCCCTGCCAGTCTTCTCCAGTGTATCTGCTTTGCA 523
QY 139 CCGCTGCCAGCCCGACACAGGTGTGAGAACCAAGTGTGTGAGGCGACTCCAGGCAC 198

XX Duda A, Finkel K, Koshy B, Parks KE;
XX WPI: 2002-154721/20.
DR P-PSDB; AAM48922.
XX Novel genetic variants of enolase 3, (beta, muscle) gene useful in
PT studying expression and function of the protein, and for screening
PT drugs to treat disorders of glycolytic pathway -
XX
PS Claim 20: Fig 1; 90pp; English.
XX
CC The present invention provides the protein, cDNA and genomic sequences of
CC a human enolase 3 (beta, muscle) isoenzyme containing a number of single
CC nucleotide polymorphisms (SNPs). The sequences can be used to identify
CC the haplotype of an individual and identify whether particular haplotypes
CC are linked to certain diseases. The present sequence is the ENO3 gene.
XX
SQ Sequence 13308 BP; 2874 A; 3775 C; 3758 G; 2901 T; 0 other;

Query Match 6.1%; Score 39.4; DB 24; Length 13308;
Best Local Similarity 45.6%; Pred. No. 2.3; Mismatches 166; Indels 0; Gaps 0;
Matches 139; Conservative 0;

QY 51 CCGAGCTCAAGATGAAGTGGGAAGGTAACAACCACTGCGTCCCTGCAAGCAGAGGC 110
DB 11006 CCGATGACACTCTTAATAATACGATCTTACAAAACCTCTGGGCCCGGACAGTGGGTC 11065
QY 111 ACTTCCGAATACCTCTCTCCCGCCGCGCTCCAGCCCAACCAAGGTGAGAAC 170
DB 11066 ACTCTTGCAATCCAGACACTTTGGAGGCTGAGCGGCGGATCACTGAGGTTAGAGT 11125
QY 171 AAGGCTGTGTGGAGCTCCAGCACTGCGCACTGCGCAACCTGCAAAATCCAT 230
DB 11126 TAGAGACCAAGCTGCGCAATGTGTGAACCCCACTCTACTAATGTACAAAAATTAC 11185
QY 231 TAGAGCACTGCGCCAGAGATGTAGAGTGTCTCAAGAGCGCTCCGAGGAGAGG 290
DB 11186 CCGGGTGTGTGACAGATGCTGTAGTCCAGCTACTGAGAGGCTAAGGAGAGATC 11245
QY 291 GACCAATCCTGTAGCTGGAAGCTGGAGAGCTCCGAAGGCCCACTACTCTCTACT 350
DB 11246 GCTTGAAACCTGGAGGAGAGGTTGACAGTGAAGCCAGATCACACCTGCACTCAGTCT 11305
QY 351 TGGTA 355
DB 11306 GGCCA 11310

RESULT 12
AAN90617
ID AAN90617 standard; cDNA; 1004 BP.
XX
AC AAN90617;
XX
DT 20-DEC-1989 (first entry)
XX
DE CD40 cDNA.
XX
KW TYROSINE SUPPRESSOR GENE; ASV LTR; human cytomegalovirus AD169 enhancer;
KW HIV box; immunoselection; immune deficiency diseases; vasculitis;
KW systemic lupus erythematosus; rheumatoid arthritis; neoplasms;
XX
PN EP330191-A.
XX
PD 30-AUG-1989.
XX
PF 23-FEB-1989; 89EP-0103127.
XX
PR 25-FEB-1988; 88US-0160416.
XX
PA (GEHO) GEN. HOSPITAL CORP.
XX

PI Seed B, Allen J, Aruffo A, Camerini D, Lauffer L, Oquendo CP,
PI Simmons D, Stamenkovic I, Stengelin S;
XX WPI: 1989-250302/35.
DR
XX
XX Rapid immunoselection cloning - used to clone genes encoding
PT cell surface antigens associated with mammalian T lymphocytes.
XX
PS Disclosure; fig. 17; 69pp; English.
XX

CC CD40 encodes a cell surface antigen involved in cell mediated immunity.
CC This DNA can be expressed in a vector which transforms COS cells. The
CC vector can isolate any protein, and clones are easy to manipulate.
XX
SQ Sequence 1004 BP; 230 A; 297 C; 276 G; 201 T; 0 other;

Query Match 6.0%; Score 39.2; DB 10; Length 1004;
Best Local Similarity 56.1%; Pred. No. 1.3; Mismatches 58; Indels 0; Gaps 0;
Matches 74; Conservative 0;

QY 79 TAACACCACTGCGTCCCTGCAAGGCGGCACTTCAGATACCTCTCCGCCAGCGC 138
DB 464 TGATACCATCTGCGAGCCCTGCGCATGCGCTTCTTCCATGTGTATCTGTTGCA 523
QY 139 CCGCTGCCAGCCCGCACACAGAGTGTGAACAAGTCTGTGTGAGCAGCTCCAGGCAC 198
DB 524 AAATGTACCTCTGGAGCAAGCTGTGAGACCAAGACTGTGTGTGCAACAGCGCAC 583
QY 199 TGCCCACTCGCA 210
DB 584 AAMCAAGACTGA 595

RESULT 13
AAQ21176
ID AAQ21176 standard; DNA; 1004 BP.
XX
AC AAQ21176;
XX
DT 21-MAY-1992 (first entry)
XX
DE Human CD40 antigen coding sequence.
XX
KW Cloning technique; cell surface antigen; immunodiagnosis; ss.
XX
OS Homo sapiens.
XX
PN WO9201049-A.
XX
PD 23-JAN-1992.
XX
PE 15-JUL-1990; 90WO-US04986.
XX
PR 13-JUL-1990; 90US-0553759.
XX
PA (GEHO-) GEN HOSPITAL CORP.
XX
PI Seed B, Aruffo A, Amlot M;
XX
DR WPI: 1992-056864/07.
XX
PT New CD53 cell surface antigen and DNA encoding it - for
PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
XX
PS Example 8; Fig 17; 160pp; English.
XX
CC Sequences coding for several human antigens were cloned using a new
CC rapid immunoselection technique (see e.g. AAQ21164). This sequence
CC encodes the CD40 antigen.
XX
SQ Sequence 1004 BP; 230 A; 298 C; 275 G; 201 T; 0 other;
XX
Query Match 6.0%; Score 39.2; DB 13; Length 1004;

Db 239 CGAGTGTACACACGTCGAGCTACATTCTGACTGCCCGCTGGCACTTAAGCCGAGCTCA 298
 QY 61 ACATGAAGTTGGGAAGGTAACACACACCTGCTCCCTGCAAGGAGGCACTT--CCAG 118
 Db 299 ACATGAAGTTGGGAAGGTAACACACCTGCTCCCTGCAAGGAGGCACTTCCAGA 358
 QY 119 AATACCTCTCCCGCAGCGCCGCTGCCAGCCCAACACAGGTGTAGAACCAAGCTGTG 178
 Db 359 ACTACCTCTCCCGCAGCGCCGCTGCCAGCCCAACACAGGTGTAGAACCAAGCTGTG 418
 QY 179 GTGGAGGGACGTCGAGCACTCCGAGTCCGACACACCACTGCAAAAATCCATTAGACCA 238
 Db 419 GTGGAGGGACGTCGAGCACTCCGAGTCCGACACACCACTGCAAAAATCCATTAGACCA 478
 QY 239 CTG 241
 Db 479 CTG 481

RESULT 9
 ABN41522
 ID ABN41522 standard; DNA; 60 BP.
 XX
 AC ABN41522;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human spliced transcript detection oligonucleotide SEQ ID NO:14270.
 XX
 KW Human; mouse; rat; splice transcript; detection; RNA transcript;
 XX splice variant; transcriptome; oligonucleotide library; ss.
 OS Homo sapiens.
 XX
 PN WO200210449-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 20-JUL-2001; 2001WO-1B01903.
 XX
 PR 28-JUL-2000; 2000US-221607P.
 XX
 PR 02-MAY-2001; 2001US-287724P.
 XX
 PA (COMP-) COMPUEN INC.
 XX
 PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
 DR WPI; 2002-257383/30.
 XX
 PT New oligonucleotide libraries comprising oligonucleotides which
 PT selectively hybridize to mRNAs transcribed from a transcription unit of
 PT a genome, useful for detecting tissue-, pathology-, and
 PT developmental-specific genes
 PS
 XX Example 1; SEQ ID 14270; 47bp; English.
 CC The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub-)transcriptome, where the
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridizing selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one or more messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterizing the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialized mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA

CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN59589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 SO Sequence 60 BP; 18 A; 15 C; 17 G; 10 T; 0 other;

Query Match 9.28; Score 60; DB 24; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.9e-06;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 ATGAAGTTGGGAAGGTAACACACACCTGCTCCCTGCAAGGAGGCACTTCCAGATA 122
 Db 1 ATGAAGTTGGGAAGGTAACACACACCTGCTCCCTGCAAGGAGGCACTTCCAGATA 60

RESULT 10
 AA51857
 ID AA51857 standard; cDNA; 845 BP.
 XX
 AC AA51857;
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE Porcine CD40 cDNA.
 XX
 KW CD40 antigen; xenograft; CD40-deficient donor; graft rejection;
 KW transgenic pig; reduced immunogenicity; immunosuppressive;
 KW gene therapy; ss.
 OS Sus scrofa.
 XX
 FH Key Location/Qualifiers
 FT CDS 9..845
 FT /*tag= a
 FT /product= CD40

WO200039294-A1.

PD 06-JUL-2000.
 XX
 PF 22-DEC-1999; 99WO-EP10332.
 XX
 PR 24-DEC-1998; 98GB-0028705.
 XX
 PR 10-FEB-1999; 99GB-0002940.
 XX

PA (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 PI Bravery C, Rushworth S, Thompson S;
 DR WPI; 2000-452392/39.

PT Ameliorating cellular graft rejection in a recipient mammal using
 PT CD40-deficient donor mammal cells, tissues or organs as a xenograft,
 PT comprises inactivation of the CD40 antigen gene with a targeting vector
 PS
 XX Claim 10; Fig 1; 23pp; English.

CC This porcine cDNA encodes CD40 antigen. A fusion protein consisting of
 CC the extracellular domain of human CD134 coupled to mouse Ig kappa light
 CC chain was able to bind to porcine CD40, indicating that direct signalling
 CC between a porcine graft and the immune cells of a human host is possible
 CC through this molecular pathway. CD40-deficient donor mammal cells,
 CC tissues or organs may be used for xenografting in order to ameliorate
 CC cellular graft rejection in the recipient mammal. In particular, porcine
 CC cells, tissues or organs are genetically modified so that the cell
 CC surface expression of CD40 antigen is reduced. Transgenic pigs may have
 CC an inactivated CD40 antigen gene and a transgene, e.g. human Decay
 CC Accelerating Factor (hDAF) gene. The CD40-deficient donor mammal cells,

XX	AC	ABL82688;
XX	ABL82688;	
DF	17-MAY-2002	(first entry)
DE	Human ovarian cancer related cDNA clone SEQ ID NO:5666.	
XX	Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.	
XX	Homo sapiens.	
OS	Homo sapiens.	
PX	MO200192581-A2.	
PN	MO200192581-A2.	
PD	06-DEC-2001.	
XX	06-DEC-2001.	
PF	29-MAY-2001; 2001WO-US17756.	
XX	29-MAY-2001; 2001WO-US17756.	
PR	26-MAY-2000; 2000US-207484P.	
XX	26-MAY-2000; 2000US-207484P.	
XX	(CORI-) CORIXA CORP.	
PI	Algate PA, Harlocker SL, Jones R;	
DR	WPI; 2002-122075/16.	
PT	Composition for therapy and diagnosis of ovarian cancer comprising	
PT	polypeptide of a ovarian tumor polypeptide, polynucleotide encoding	
PT	polypeptide, antibody specific to polypeptide or T cell expressing	
PT	polypeptide -	
PS	Claim 1; SEQ ID 5666; 489pp: English.	
XX		
CC	The present invention describes a composition (I) comprising: carriers	
CC	and immunostimulants; and a polypeptide (II) of a ovarian tumour	
CC	polypeptide encoded by a polynucleotide (III) having a cDNA sequence	
CC	(S1) from the 10912 nucleotide sequences as given in ABL77023 to	
CC	ABL87934, (III) encoding (II) having a sequence (S2), a T cell	
CC	population of (II), or antigen presenting cells that express (II).	
CC	(C) has cytostatic activity. An oligonucleotide (IV) that hybridises to	
CC	(S1) can be used for detecting ovarian cancer in a patient's biological	
CC	sample preferably serum or ovarian tissue. The method comprises	
CC	contacting a biological sample from a patient with (IV), detecting the	
CC	amount of polynucleotide hybridising to (IV) and comparing the amount to	
CC	a predetermined cutoff value and thereby detecting ovarian cancer in the	
CC	patient, where the amount of polynucleotide hybridising to (IV) is	
CC	detected preferably by polymerase chain reaction (PCR). (I) comprising	
CC	(III) and/or (II) is useful for stimulating and/or expanding T cells	
CC	specific for an ovarian tumour protein comprising contacting T cells	
CC	with (III) or (II). (III) is useful in design and preparation of	
CC	ribozyme molecules for inhibiting expression of the tumour polypeptides	
CC	and proteins in tumour cells; and to isolate a full length gene from a	
CC	suitable library e.g., a tumour cDNA library using well known	
CC	techniques.	
SQ	Sequence 289 BP; 49 A; 75 C; 93 G; 72 T; 0 other:	
XX		
XX		
OY	Query Match	38.7%; Score 251.2; DB 24; Length 289;
OY	Best Local Similarity	98.5%; Pred. No.1e-56; Mismatches 3; Indels 1; Gaps
OY	Matches 264; Conservative 0;	
Db	379 AGATGTTTCCCAGTATTCACCTGGGCGTCCCGCAGGCCCAAGTTTGGAGGCAGGGGTGCC 438	
Db	289 AGATGTTTCCCAGTATTCACCTGGGCGTCCCGCAGGCCCAAGTTTGGAGGCAGGGGTGCC 230	
OY	439 GCACAGCAGA-GTCTCTTGGAACCTGACACAGGAGACCAGCACTTGGAAACCGGGAGCAGA 497	
Db	229 GCACAGCAGAATGTCTCTTGGAACCTGACACAGGAGACCAGCACTTGGAAACCGGGAGCAGA 170	
OY	498 .GCACAGTGGCCACAGGTGACCAATGGCATTCATTGACACGGGGGGTCTATGACTTCACTG 557	
Db	169 GCCAGGTGGCCACAGGTGACCAATGGCATTCATTGACACGGGGGGTCTATGACTTCACTG 110	
OY	558 GCACATCTACATCTACATGACCAAGTACTGGGGGGACCAACCGGGGTCTTGAGAACCTGC 617	

Db	109	GCACATCTCATCTACATGACACATCTACTGCGGGGACACCGGGTCTCGAGACCTCC	50
QY	618	CAGCTACCGCGAACCTTCATATCCCAT 645	
Db	49	CAGCTACCGCGAACCTTCATCTCCCT 22	
RESULT 8			
XX	AB181903	standard; cDNA; 481 BP.	
AC	AB181903;		
DT	17-MAY-2002	(first entry)	
DE	Human ovarian cancer related cDNA clone SEQ ID NO:4881.		
KW	Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.		
OS	Homo sapiens.		
PN	WO200192581-A2.		
PD	06-DEC-2001.		
PF	29-MAY-2001; 2001WO-US17756.		
PR	26-MAY-2000; 2000US-207484P.		
PA	(CORI-) CORIXA CORP.		
PI	Algate PA; Harlocker SL; Jones R;		
DR	WPI: 2002-122075/16.		
XX	Composition for therapy and diagnosis of ovarian cancer comprising		
PT	polypeptide of a ovarian tumor polypeptide, polynucleotide encoding		
PT	polypeptide, antibody specific to polypeptide or T cell expressing		
XX	polypeptide		
PS	Claim 1; SEQ ID 4881; 489pp; English.		
XX	The present invention describes a composition (I) comprising: carriers		
CC	and immunostimulants; and a polypeptide (II) of a ovarian tumour		
CC	polypeptide encoded by a polynucleotide (III) having a cDNA sequence		
CC	(S1) from the 10912 nucleotide sequences as given in ABL77023 to		
CC	ABL87934, (III) encoding (II) having a sequence (S2), a T cell		
CC	population of (II), or antigen presenting cells that express (II).		
CC	(I) has cytostatic activity. An oligonucleotide (IV) that hybridises to		
CC	(SI) can be used for detecting ovarian cancer in a patient's biological		
CC	sample preferably serum or ovarian tissue. The method comprises		
CC	contacting a biological sample from a patient with (IV), detecting the		
CC	amount of polynucleotide hybridising to (IV) and comparing the amount to		
CC	a predetermined cutoff value and thereby detecting ovarian cancer in the		
CC	patient, where the amount of polynucleotide hybridising to (IV) is		
CC	detected preferably by polymerase chain reaction (PCR). (I) comprising		
CC	(III) and/or (II) is useful for stimulating and/or expanding T cells		
CC	specific for an ovarian tumour protein comprising contacting T cells		
CC	with (III) or (II). (III) is useful in design and preparation of		
CC	ribozyme molecules for inhibiting expression of the tumour polypeptides		
CC	and proteins in tumour cells; and to isolate a full length gene from a		
CC	suitable library e.g., a tumour cDNA library using well known		
CC	techniques.		
SO	Sequence 481 BP; 113 A; 161 C; 123 G; 84 T; 0 other;		
QY	Query Match	32.1%; Score 208.2; DB 24; Length 481;	
	Best Local Similarity	93.8%; Pred. No. 2.6e-45;	
	Matches 228; Conservative 0; Mismatches 13; Indels 2; Gaps 14;		
1	CGAGTGCACACCTCGAGCTCTTTCGACGCGCGCGCTGCAGTGAAGCCGAGTCAA 60		

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pt_sequences.

XX Sequence 1594 BP; 391 A; 488 C; 434 G; 280 T; 1 other;

Query Match 72.4%; Score 470; DB 24; Length 1594;

Best Local Similarity 84.4%; Pred. No. 2,8e-114;
 Matches 588; Conservative 0; Mismatches 1; Indels 108; Gaps 1;

```

QY 61 AGATGAAGTTGGAGGAGTACACCACTGCTCCCTGCAGAGGAGGACCTTCCAGAA 120
   |||||||
DB 91 AGATGAAGTTGGAGGAGTACACCACTGCTCCCTGCAGAGGAGGACCTTCCAGAA 150
   |||||||
QY 121 TACCTCTCCCTCCAGGCGCGTGCAGCGCCACACAGGTGTGAGAACCAAGTCTGTT 180
   |||||||
DB 151 TACCTCTCCCTCCAGGCGCGTGCAGCGCCACACAGGTGTGAGAACCAAGTCTGTT 210
   |||||||
QY 181 GGAGGAGAGCTCCAGGAGCTCCAGTCCGACACACACCTGCAAAAATTCATTAGAGCACT 240
   |||||||
DB 211 GGAGGAGAGCTCCAGGAGCTCCAGTCCGACACACACCTGCAAAAATTCATTAGAGCACT 270
   |||||||
QY 241 GCGCCAGAGAGATGCA----- 256
   |||||||
DB 271 GCGCCAGAGAGATGCAAGAACATGCTGATGCTGCGCTTCTGCTGCACTGCGCTTCTT 330
   |||||||
QY 257 ----- 256
   |||||||
DB 331 TCTGCTCTTCCACCGTCTTCTCTGACATCTGGAAAGGACACCTTCTCTGACAGAA 380
   |||||||
QY 257 ----GGATGCTCTCTCAAGAGGCGTCCGACGAGGAGAGGAGACCAATCTGTAGCTGAGAG 312
   |||||||
DB 391 ACTGGATCTGCTCTCAAGAGGCGTCCGACGAGGAGAGGAGACCAATCTGTAGCTGAGAG 450
   |||||||
QY 313 CTGGGAGCGCTCCGAAGCGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
   |||||||
DB 451 CTGGGAGCGCTCCGAAGCGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
   |||||||
QY 373 TTCTGGAGATGTTTCCCAATATCATCTGAGTCTCCCGGACGCGCCAGTTTGGAGGAGG 432
   |||||||
DB 511 TTCTGGAGATGTTTCCCAATATCATCTGAGTCTCCCGGACGCGCCAGTTTGGAGGAGG 570
   |||||||
QY 433 GGTGCGGACACAGCAGAGTCTCTGACCTTGACAGGAGGAGCGGAGTTGGAACCGGGGA 492
   |||||||
DB 571 GGTGCGGACACAGCAGAGTCTCTGACCTTGACAGGAGGAGCGGAGTTGGAACCGGGGA 630
   |||||||
QY 493 GCAGAGCCAGGTGCGCCAGGATGACATGATGATGATGATGATGATGATGATGATGATGAT 552
   |||||||
DB 631 GCAGAGCCAGGTGCGCCAGGATGACATGATGATGATGATGATGATGATGATGATGATGAT 690
   |||||||
QY 553 CACTGGCAGATCTACATCTACATGAGACCAAGTACTGGGGGAGACACCGGGTCTTGAGGA 612
   |||||||
DB 691 CACTGGCAGATCTACATCTACATGAGACCAAGTACTGGGGGAGACACCGGGTCTTGAGGA 750
   |||||||
QY 613 CTTCCAGCTACCCCGAACCCTGCATCCCATTTCCC 649
   |||||||
DB 751 CTTCCAGCTACCCCGAACCCTGCATCCCATTTCCC 787
   |||||||

```

RESULT 6
 ABL82623
 ID ABL82623 standard; cDNA; 281 BP.

AC ABL82623;

XX 17-MAY-2002 (first entry)

DE Human ovarian cancer related cDNA clone SEQ ID NO:5601.

XX Human; ovarian cancer; ovarian tumour; cytosolic; gene; ss.

OS Homo sapiens.

PN WO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US17756.

XX 26-MAY-2000; 2000US-207484P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.

PT Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PS polypeptide
 XX Claim 1; SEQ ID 5601; 489pp; English.

CC The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 CC (S1) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

XX Sequence 281 BP; 58 A; 94 C; 77 G; 52 T; 0 other;

Query Match 39.1%; Score 253.8; DB 24; Length 281;

Best Local Similarity 97.1%; Pred. No. 2e-57;
 Matches 269; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

```

QY 368 CCCATTCTGAGATGTTTCCCAATATCATCTGAGTCTCCCGCAGGCC-CAGTTTGGGA 426
   |||||||
DB 5 CCACTTCTGAGATGTTTCCCAATATCATCTGAGTCTCCCGCAGGCCAGTTTGGGA 64
   |||||||
QY 427 GCGAGGGGTGCGCCAGAGAGTCTCTGACCTTGACAGGAGGAGCGGAGTTGGAAC 486
   |||||||
DB 65 GCGAGGGGTGCGCCAGAGAGTCTCTGACCTTGACAGGAGGAGCGGAGTTGGAAC 124
   |||||||
QY 487 CCGGAGACAGAGCAGGTGCGCCAGGATGACATGATGATGATGATGATGATGATGATGATGAT 546
   |||||||
DB 125 CCGGAGACAGAGCAGGTGCGCCAGGATGACATGATGATGATGATGATGATGATGATGATGAT 184
   |||||||
QY 547 GACTATCATCTGGAACATCTACATCTACATGAGACCAAGTACTGGGGGAGACACCGGGTCC 606
   |||||||
DB 185 GACTATCATCTGGAACATCTACATCTACATGAGACCAAGTACTGGGGGAGACACCGGGTCA 244
   |||||||
QY 607 TGGAGACCTCCCGAACCCTGCATCCCATTTCCC 643
   |||||||
DB 245 TGGAGACCTCCCGAACCCTGCATCCCATTTCCC 281
   |||||||

```

RESULT 7
 ABL82688/C
 ID ABL82688 standard; cDNA; 289 BP.

CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 81.8%; Score 531; DB 24; Length 2136;

Best Local Similarity 85.7%; Pred. No. 2,4e-130;

Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

QY 1 CGAGGTACACACTGCGAGCTACTTCTGACTGCTCCCGCTGCGACTGAGCCGAGCTCAA 60
 DB 579 CGAGGTACACACTGCGAGCTACTTCTGACTGCTCCCGCTGCGACTGAGCCGAGCTCAA 638
 QY 61 AGATGAGTTGGGAGGAGGTAACACACCTGCTCCCTTCGAGGAGGCGACTTCCAGAA 120
 DB 639 AGATGAGTTGGGAGGAGGTAACACACCTGCTCCCTTCGAGGAGGCGACTTCCAGAA 698
 QY 121 TACCTCTCCCGCGAGGCGCGCTGCGAGCCGCGACACAGGTGAGAACCAAGTCTGCT 180
 DB 699 TACCTCTCCCGCGAGGCGCGCTGCGAGCCGCGACACAGGTGAGAACCAAGTCTGCT 758
 QY 181 GGAGGAGCTCCAGGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCT 240
 DB 759 GGAGGAGCTCCAGGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCTGCGAGCT 818
 QY 241 GCGCCGAGAGATGTCA----- 256
 DB 819 GCGCCGAGAGATGTCAAGAACCATGCTGATGCTGCGCTTCTGCTGCTGCTGCTGCTGCT 878
 QY 257 ----- 256
 DB 879 TCTGCTCTTGGCAGCGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 938
 QY 257 ----- 256
 DB 939 ACTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 998
 QY 313 CTGGGAGCTCCGAGGAGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
 DB 999 CTGGGAGCTCCGAGGAGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1058
 QY 373 TTTGAGAGATGTTTCCCAAGTATCCACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432
 DB 1059 TTTGAGAGATGTTTCCCAAGTATCCACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1118
 QY 433 GGTGCGGAGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 492
 DB 1119 GGTGCGGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1178
 QY 493 GCAGAGGAGGTTGGGCGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
 DB 1179 GCAGAGGAGGTTGGGCGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1238
 QY 553 CACTGGCAACATCTACATCTACATGAGACAGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 612
 DB 1239 CACTGGCAACATCTACATCTACATGAGACAGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1298
 QY 613 CCTCCAGCTACCCCGGAGCTCCATACCCATTCCTCC 649
 DB 1299 CCTCCAGCTACCCCGGAGCTCCATACCCATTCCTCC 1335

RESULT 5
 AB055003
 ID AB055003 standard; cDNA; 1594 BP.
 XX
 AC AB055003;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Human ovarian antigen HSABJ44 cDNA, SEQ ID NO:883.
 XX
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynecological; reproductive; chromosome 12p13;
 KW gene; ss.
 KW
 KW Homo sapiens.
 OS
 PN WO200200677-A1.
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001WO-US18569.
 PR 07-JUN-2000; 2000US-209467P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 DR WPI; 2002-147878/19.
 XX
 PT P-PSDB; ABP41926.
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 1; SEQ ID NO 883; 2922bp; English.
 CC
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (AB054131-AB056305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovarian and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention.


```

QY      1 CGAGTGTACACACTGCGAGCTACTTCTTGACMGCCCGCTGGCACTGAAGCCGAGCTCAA 60
        |||
Db      579 CGAGTGTACACACTGCGAGCTACTTCTTGACMGCCCGCTGGCACTGAAGCCGAGCTCAA 638
QY      61 AGATGAAGTTGGGAAGGGTAACACACACTGCTGCTCCCTGCAAGGAGGCACTTCCAGAA 120
        |||
Db      639 AGATGAAGTTGGGAAGGGTAACACACACTGCTGCTCCCTGCAAGGAGGCACTTCCAGAA 698
QY      121 TACCTCTCCCGCCAGCGCCGCTGCCAGGCCCAACACAGGTGTGAACCAAGGCTGTGT 180
        |||
Db      699 TACCTCTCCCGCCAGCGCCGCTGCCAGGCCCAACAGGTGTGAACCAAGGCTGTGT 758
QY      181 GAGGAGAGCTCCAGGCACTGCGCAGTCCAGACACAACTGCAAAAAATTCATTAGAGCCACT 240
        |||
Db      759 GAGGAGAGCTCCAGGCACTGCGCAGTCCAGACACAACTGCAAAAAATTCATTAGAGCCACT 818
QY      241 GCCCCAGAGATGTCA----- 256
        |||
Db      819 GCCCCAGAGATGTCAAGAACATGCTGATGCTGCCCTTCTGCTGACCACTGCGCTTCTT 878
        |||
QY      257 ----- 256
        |||
Db      879 TCTGCTCTTCCACACCGCTCTTCTCTGCACTGTGGAAGACCACTTCTCTCTGAGGAA 938
QY      257 -----GATTCGCTCTCTCAAGAGGGGTCGCGAGGAGAGGACCCATCTCTGTAGTGGAG 312
        |||
Db      939 ACTGAGATCTGCTCTCTCAAGAGGGGTCGCGAGGAGAGGACCCATCTCTGTAGTGGAG 998
QY      313 CTGGAGACCTCCGCAAGGCGCATCATCTCTGACTTGTGACAGCCACTGCTACCCAT 372
        |||
Db      999 CTGGAGACCTCCGCAAGGCGCATCATCTCTGACTTGTGACAGCCACTGCTACCCAT 1058
QY      373 TTCTGAGATGTTTCCCACTATCTGAGGCTCCCGCAGGCCCACTTTTGAAGCAGG 432
        |||
Db      1059 TTCTGAGATGTTTCCCACTATCTGAGGCTCCCGCAGGCCCACTTTTGAAGCAGG 1118
QY      433 GGTGCTCCGCAAGAGAGTCTCTGAGACCTGACAGGAGCCGAGTTGGAACCCGAGGA 492
        |||
Db      1119 GGTGCTCCGCAAGAGAGTCTCTGAGACCTGACAGGAGCCGAGTTGGAACCCGAGGA 1178
QY      493 GCAGAGCAGAGTGGCCAGGCTACCAATGAGATTCATGTACCGGCGGCTATGACTAT 552
        |||
Db      1179 GCAGAGCAGAGTGGCCAGGCTACCAATGAGATTCATGTACCGGCGGCTATGACTAT 1238
QY      553 CACTGGCAACATCTACATCTACAAATGACAGTACTGGGGGGACACCGGGTCTCTGAGA 612
        |||
Db      1239 CACTGGCAACATCTACATCTACAAATGACAGTACTGGGGGGACACCGGGTCTCTGAGA 1298
QY      613 CCTCCAGCTACCCCGGCAACCTCATACCCCATCC 649
        |||
Db      1299 CCTCCAGCTACCCCGGCAACCTCATACCCCATCC 1335

RESULT 3
ABL64078
ID      ABL64078 standard; DNA; 2136 BP.
XX      ABL64078;
AC      15-MAY-2002 (first entry)
DE      Breast cancer related gene sequence SEQ ID NO:2415.
XX      15-MAY-2002 (first entry)
KW      Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
        stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
        cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
        gene; ds.
XX      Homo sapiens.
OS      Homo sapiens.
PN      WO200194629-A2.
XX      13-DEC-2001.
PD

```

```

XX      30-MAY-2001; 2001WO-US10838.
PF      05-JUN-2000; 2000US-209473P.
XX      05-JUN-2000; 2000US-209531P.
PR      18-SEP-2000; 2000US-233133P.
PR      18-SEP-2000; 2000US-233617P.
PR      20-SEP-2000; 2000US-234009P.
PR      20-SEP-2000; 2000US-234034P.
PR      20-SEP-2000; 2000US-234052P.
PR      22-SEP-2000; 2000US-234509P.
PR      22-SEP-2000; 2000US-234567P.
PR      25-SEP-2000; 2000US-234923P.
PR      25-SEP-2000; 2000US-234924P.
PR      25-SEP-2000; 2000US-235077P.
PR      25-SEP-2000; 2000US-235082P.
PR      25-SEP-2000; 2000US-235134P.
PR      25-SEP-2000; 2000US-235280P.
PR      26-SEP-2000; 2000US-235637P.
PR      26-SEP-2000; 2000US-235638P.
PR      26-SEP-2000; 2000US-235711P.
PR      27-SEP-2000; 2000US-235720P.
PR      27-SEP-2000; 2000US-235840P.
PR      27-SEP-2000; 2000US-235863P.
PR      28-SEP-2000; 2000US-236028P.
PR      28-SEP-2000; 2000US-236032P.
PR      28-SEP-2000; 2000US-236033P.
PR      28-SEP-2000; 2000US-236034P.
PR      28-SEP-2000; 2000US-236109P.
PR      28-SEP-2000; 2000US-236111P.
PR      29-SEP-2000; 2000US-236842P.
PR      29-SEP-2000; 2000US-236891P.
PR      02-OCT-2000; 2000US-237172P.
PR      02-OCT-2000; 2000US-237173P.
PR      02-OCT-2000; 2000US-237278P.
PR      02-OCT-2000; 2000US-237294P.
PR      02-OCT-2000; 2000US-237295P.
PR      02-OCT-2000; 2000US-237316P.
PR      03-OCT-2000; 2000US-237425P.
PR      03-OCT-2000; 2000US-237598P.
PR      03-OCT-2000; 2000US-237604P.
PR      03-OCT-2000; 2000US-237606P.
PR      03-OCT-2000; 2000US-237608P.
PR      01-NOV-2000; 2000US-244867P.
PR      01-NOV-2000; 2000US-245084P.
XX      (AVAL-) AVALON PHARM.
PA      Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI      Soppet DR, Weaver Z;
PI      WPI; 2002-188264/24.
XX      Screening for anti-neoplastic agent involves exposing cells to a
        chemical agent to be tested for anti-neoplastic activity, and
        determining a change in expression of a gene of a signature gene set
        Claim 1; SEQ ID 2415; 44pp; English.
XX      The present invention describes a method (M1) for screening for an
        anti-neoplastic agent. The method involves exposing cells to a chemical
        agent to be tested for anti-neoplastic activity, determining a change in
        expression of at least one gene (I) of a signature gene set, where (I)
        comprises a sequence (S) selected from 8447 sequences (given in ABL61664
        to ABL70110), or is at least 95% identical to (S), where a change in
        expression is indicative of anti-neoplastic activity. (I) has cytostatic
        activity and can be used in gene therapy. M1 can be used for screening
        an anti-neoplastic agent, and can be used for producing a product which
        is the data collected with respect to the anti-neoplastic agent as a
        result of M1, and the data is sufficient to convey the chemical
        structure and/or properties of the agent. M1 can be used in the
        treatment of cancer such as colon, breast, stomach, lung, thyroid,
        oesophageal, ovarian, kidney, prostate or pancreatic cancer,

```

PT Involves detecting the level of expression of two or more genes in a
PT Liver tissue sample
XX Claim 1: SEQ ID NO 2135; 298bp; English.
PS
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pot_sequences.
XX
SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 81.8%; Score 531; DB 24; Length 2136;
Best Local Similarity 85.7%; Pred. No. 2.4e-130;
Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

QY 1 CGAGTGACACTGCGAGCTACTTCTGACTCCCGCCCTGGACATGAGCCGAGCTCAA 60
DB 579 CGAGTGACACTGCGAGCTACTTCTGACTCCCGCCCTGGACATGAGCCGAGCTCAA 638
QY 61 AGATGAAGTTGGAGGATTAACAACCACTGCTCCCTGCAAGGAGGAGCACTTCCAGAA 120
DB 639 AGATGAAGTTGGAGGATTAACAACCACTGCTCCCTGCAAGGAGGAGCACTTCCAGAA 638
QY 121 TACCTCTCTCCCGAGGCGCGCTGCGACCCGACACAGAGTGGAGAACCAAGTCTGGT 180
DB 699 TACCTCTCTCCCGAGGCGCGCTGCGACCCGACACAGAGTGGAGAACCAAGTCTGGT 758
QY 181 GGAGGAGCTCCAGGAGCACTGCTCCAGTCCGACACAACTGCAAAAATCATTTAGAGCCACT 240
DB 759 GGAGGAGCTCCAGGAGCACTGCTCCAGTCCGACACAACTGCAAAAATCATTTAGAGCCACT 818
QY 241 GCGCCCGAGAGATGCA----- 256
DB 819 GCGCCCGAGAGATGCAAGAACATGCTGATGCGCCGTTCTGCTGCACTGGGCTTCTT 878
QY 257 ----- 256
DB 879 TCTGCTCTTTCGCAACCGTCTTCTCTGATCTGAGAGAGCCACCTTCTCTGCAAGAA 938
QY 257 -----GATCGCTGCTCAAGAGGCGTCCGAGGAGAGAGAGCCCAATCTCTGAGCTGGAAG 312
DB 939 ACTGCGATCTGCTGCTCAAGAGGCGTCCGAGGAGAGAGAGCCCAATCTCTGAGCTGGAAG 998
QY 313 CTGGAGACCTTCGAAGGCGCCATCATCTTCCCTGACTTGTACAGCCACTGCTACCCAT 372
DB 999 CTGGAGACCTTCGAAGGCGCCATCATCTTCCCTGACTTGTACAGCCACTGCTACCCAT 1058
QY 373 TTCTGGAGATTTTCCCAATATTCATCTGGGCTTCCCGAGCCCAATTTTGGAGGAGG 432
DB 1059 TTCTGGAGATTTTCCCAATATTCATCTGGGCTTCCCGAGCCCAATTTTGGAGGAGG 1118
QY 433 GGTGCGCCACAGACAGAGTCTCTGACCTGACAGGAGAGCCGAGTGGAAACCCGGGGA 492
DB 1119 GGTGCGCCACAGACAGAGTCTCTGACCTGACAGGAGAGCCGAGTGGAAACCCGGGGA 1178
QY 493 GCAGAGCCAGAGTGGCCAGCGGTACCAATGAGCATGATGATCGCGGGGTCTATGACTAT 552
DB 1179 GCAGAGCCAGAGTGGCCAGCGGTACCAATGAGCATGATGATCGCGGGGTCTATGACTAT 1238
QY 553 CACGCGCAATATCTACTATCAATGAGCATGATGAGGAGGAGCCAGCGGGGTCTCTGAGA 612
DB 1238 CACGCGCAATATCTACTATCAATGAGCATGATGAGGAGGAGCCAGCGGGGTCTCTGAGA 612

DB 1239 CACTGGCAACATCTACTATCTACATGAGACCACTGCGGGGAGCCAGCGGGTCTTGAGA 1298
QY 613 CCTCCAGCTACCCCGCAACCTCCATACCCCATTTCC 649
DB 1299 CCTCCAGCTACCCCGCAACCTCCATACCCCATTTCC 1335

RESULT 2
ABK64502
ID ABK64502 standard; DNA; 2136 BP.
XX
AC ABK64502;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human benign prostatic hyperplasia gene #397.
XX
KW Human: benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200212440-A2.
XX
PD 14-FEB-2002.
XX
PF 07-AUG-2001; 2001WO-US24708.
XX
PR 07-AUG-2000; 2000US-22323P.
XX
PR 05-JUN-2001; 2001US-0873319.
XX
PA (GENE-) GENE LOGIC INC.
XX
PA (NISB) JAPAN TOBACCO INC.
XX
PI Munger WE, Kulkarl P, Getzenberg RH, Waga I, Yamamoto J;
XX
DR WPI; 2002-257476/30.
XX
XX
PT Identifying drugs for and diagnosing benign prostatic hyperplasia, by
PT detecting expression levels of one or more genes in prostate cells from
PT patient that are differentially regulated compared to normal prostate
PT cells -
XX
XX
PS Disclosure; Page 229-230; 444pp; English.
XX
CC The invention relates to a method of diagnosing (I) the onset or
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC or identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles.
CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
CC useful for identifying an agent that modulates the onset or progression
CC of BPH. The methods are useful to present information identifying
CC the expression level in a tissue or cells, by comparing the expression
CC level of genes given in the specification in the tissue or cells to the
CC level of expression of gene in the database, and displaying the
CC expression levels of at least one gene in the tissue or cell sample
CC compared to the expression level in BPH. Agents using (II) are useful for
CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
CC benign prostatic hyperplasia gene sequences of the invention.
XX
SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;.

Query Match 81.8%; Score 531; DB 24; Length 2136;
Best Local Similarity 85.7%; Pred. No. 2.4e-130;
Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

Run on: April 15, 2003, 13:05:15 ; Search time 112.963 Seconds
(without alignments)

Perfect score: 649
Sequence: 1 cgagtgtacacactgagc.....aacctccatacccatccc 649

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

```
Database : N_Geneseq_101002:*
1. /STDS2/acqdata/geneseq/geneseqn-emb1/NA1980.DAT:*
```

1:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT *
2:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT *
3:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT *
4:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT *
5:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT *
6:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT *
7:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT *
8:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT *
9:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT *
10:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT *
11:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT *
12:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT *
13:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT *
14:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT *
15:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT *
16:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT *
17:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT *
18:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT *
19:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT *
20:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT *
21:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT *
22:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT *
23:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT *
24:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	531	81.8	2136	24	ABN95637	Gene #2135 used to
2	531	81.8	2136	24	ABK64502	Human dengjin prot
3	531	81.8	2136	24	ABL64078	Breast cancer rela
4	531	81.8	2136	24	ABL64384	Stomach cancer rel
5	470	72.4	1594	24	ABO55003	Human ovarian anti
6	253.8	39.1	281	24	ABL82523	Human ovarian canc
7	251.2	38.9	289	24	ABL82688	Human ovarian canc
8	208.2	32.1	481	24	ABL81903	Human ovarian canc
9	60	9.2	60	24	ABN41522	Human spliced tran

10	40.8	6.3	845	21	AA51857	Porcine CD40 CDNA.
11	39.4	6.1	13308	24	AAAN96531	Human enolase 3 (be
12	39.2	6.0	10034	20	AAAN98617	CD40 CDNA. AAN90
13	39.2	6.0	1004	13	AAO21176	Human CD40 antigen
14	39.2	6.0	1004	14	AAO47341	CD40 coding sequen
15	39.2	6.0	1004	17	AAFT14706	Human CD40 antigen
16	39.2	6.0	1004	19	AAV63454	Human CD40 antigen
17	39.2	6.0	1004	20	AAV20296	CD40 CDNA. uniden
18	39.2	6.0	1004	20	AA423432	Human CD40 DNA. H
19	39.2	6.0	1004	20	AA424936	Human CD40 encodin
20	39.2	6.0	1004	21	AAV81198	Human CD40 antigen
21	39.2	6.0	1004	21	AAV50590	Human cell surface
22	39.2	6.0	1004	22	AA477769	Human CD40 nucleot
23	39.2	6.0	1004	22	AA503184	Human lymphocyte c
24	39.2	6.0	1004	24	AA147328	CD40 B-cell antigen
25	39	6.0	11244	22	AAK74606	Human immune/haema
26	39	6.0	15296	22	AAK74605	Human immune/haema
27	38.8	6.0	15044	22	AA536290	Human cardiovascular
28	38.6	5.9	15046	22	AA536291	Human cardiovascular
29	38.6	5.9	765	21	AAZ55529	Human cardiovascular
30	38.6	5.9	765	21	AAZ55530	Canine mature CD40
31	38.6	5.9	822	21	AAZ55528	Canine mature CD40
32	38.6	5.9	822	21	AAZ55528	Canine CD40 CDNA c
33	38.6	5.9	1425	21	AAZ55525	Canine CD40 CDNA c
34	38.6	5.9	1425	21	AAZ55526	Canine CD40 CDNA.
35	37.4	5.8	6536	19	AAV04471	Canine CD40 CDNA c
36	37.4	5.8	6536	19	AAV99118	Nduarelia beta-11
37	36.6	5.6	333	20	AAV72224	Nduarelia beta-11
38	36.6	5.6	459	20	AAAG6451	Human anti-Grp78/I
39	36.4	5.6	459	20	AAAG6452	Human TNFR superfa
40	36.4	5.6	767	21	AAAX2619	Human TNFR soluble
41	36.4	5.6	813	20	AAZ23377	Human APOE DNA. H
42	36.4	5.6	813	21	AAV75999	Human mFLN #1 nu
43	36.4	5.6	813	21	AAAH8730	DNA encoding a matD
44	36.4	5.6	813	21	AAAS1077	Human FAS ligand 1
45	36.4	5.6	813	22	AAAD07380	Mature human FLN1

RESULT 1	
ABN95637	
ID	ABN95637 standard; DNA; 2136 BP.
XX	
AC	ABN95637;
XX	
DT	13-AUG-2002 (first entry)
XX	
DE	Gene #2135 used to diagnose liver cancer.
XX	
KW	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KM	metastatic liver tumour; cytostatic; expression profile; disease strate;
XX	disease progression; drug toxicity; drug efficacy; drug metabolism.
OS	
XX	Homo sapiens.
PN	
XX	WO200229103-A2.
PD	
XX	11-APR-2002.
PF	
XX	02-OCT-2001; 2001WO-US30589.
PR	
XX	02-OCT-2000; 2000US-237054P.
PA	
XX	(GENE-) GENE LOGIC INC.
PI	
XX	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
DR	
XX	WPI; 2002-426119/45.
PT	
PT	diagnosing and detecting the progression of liver cancer,
PT	hepatocellular carcinoma or metastatic liver tumor in a patient,

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Matches 474; Conservative 0; Mismatches 53; Indels 7; Gaps 5;

```
QY 76 CCCGGCCCTGGGGTGCACATCGGCCCTGAGTCCCGTCCAGAGCTCTGGGCTCGGGCAGCC 135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CCCGGCCCTGGGGTGCACATCGGCCCTGAGTCCCGTCCAGAGCTCTGGGCTCGGGCAGCC 60

QY 136 GCCCGCACCCGCTGCGCCAGGAGCTGCGGCTCTCCCTCCAGGCCCCCAGCTGCT 195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 G-CGCCANCGCTGCCAGGAGCTGCGGCTCTCTTTTTCCTCGGCTCTTTTTC 119

QY 196 GCGCGCTGCGCGAGTGGCGCCCATG--CTCCCTGCTTGGGCGACCTCTGCCCGGCT 253
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 120 CCCTCTGGCGAGCTTGGCGGCTCATGGCTCTGCTCTTGGGAGCTCTGCTCCCGGCT 179

QY 254 GGCCTGCGGCGCTCTGCTGCTGGGCTCTTTCGGGCTCTCGGAGATCGACGCCAGGC 313
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 GGCCTGCGGCG-CTCTGGTGTGGGCTCTTCGGGCTCTCTGCGAGCATCGCAGCCAG-- 236

QY 314 GGTGCTCCATATGCGCGGAGAACCAAGACCTGAGGGACCGAGAAAGAAATACTATGA 373
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 237 CGGTCTCATATGCGTGGAGAACCAAGACCTGAGGGACCGAGAAAGAAATACTATGA 296

QY 374 GCCCAGCAGCGCATCTGCTGCTCCGCTGCCGCGCAGGCACTATGCTCAGCTAAATG 433
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 297 GCCCAGCAGCGCATCTGCTGCTCCGCTGCCGCGCA-GCAGCTATGCTCAGCTAAATG 355

QY 434 TAGCCGATCCGGAGACAGATTGTGCCACATGTGCGAGAAATTCCTACAAAGAGACTG 493
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 356 TAGCCGATCCGGAGACAGATTGTGCCACATGTGCGAGAAATTCCTACAAAGAGACTG 415

QY 494 GAACTACCTGACATGCGCGAGCTGCGCGCCCTGTGACCCAGTGGGCTCGAGGA 553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 416 GAACTACCTGACATGCGCGAGCTGCTGCGCGCCCTGTGACCCAGTGGGCTCGAGGA 475

QY 554 GATTGCCCTCTGCACAAGCAAAACGAAAGACCCAGTGGCGCTGCCAGCGGGGAT 607
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 476 GATTGCCCTCTGCACAAGCAAAACGAAAGACCCAGTGGCGCTGCCAGCGGGGAT 529
```

Search completed: April 15, 2003, 18:45:36
Job time : 150.101 secs

Db 1 CTGAGCTCCGCGCATTGGAGAGCCCTGAGGCCCTGGCCCTCCCGCCCTGGGGGTGCA 60
QY 93 CATAGGCCCTGAGTCCGCTCCAGAGCTCTGGGCTCGGGCAGCCGCGCCAGCCGCTGCCCA 152
Db 61 CATGGGCCCTGAGTCCGCTCCAGAGCTCTGGGCTCGGGCAGCCGCGCCAGCCGCTGCCCA 120
QY 153 GAGAGTGGGGCTCTGCTGCTCTCCAGAGCCGCCACGCTTGCGCCGCTGGCGAGTGT 212
Db 121 GGAGCTGGGGCTCTGCTGCTCTCCAGAGCCGCCACGCTTGCGCCGCTGGCGAGTGT 180
QY 213 GCCGCGATGCTCTGCTGCTCTCCAGAGCCGCCACGCTTGCGCCGCTGGCGAGTGT 272
Db 181 GCCGCGATGCTCTGCTGCTCTCCAGAGCCGCCACGCTTGCGCCGCTGGCGAGTGT 240
QY 273 CTGGGCCCTCTGCGGCTCTCTGAGAGATCGAGCCCGAGGGGCTCTGATGAGTCCGCTG 332
Db 241 CTGGGCCCTCTGCGGCTCTCTGAGAGATCGAGCCCGAGGGGCTCTGATGAGTCCGCTG 300
QY 333 GAGAACCGAGCTCGAGGGAGCCAGAGAAAGAAATCTATGAGCCCGACCGCATCTGC 392
Db 301 GAGAACCGAGCTCGAGGGAGCCAGAGAAAGAAATCTATGAGCCCGACCGCATCTGC 360
QY 393 TGTCCCGGCTCGGCGCGAGAGCTATGCTCAGCTAATGATGAGCCGATCCGGAGACA 452
Db 361 TGTCCCGGCTCGGCGCGAGAGCTATGCTCAGCTAATGATGAGCCGATCCGGAGACA 420
QY 453 GTTGTGCGCATGTGCGGAGAAATCTCACAAGCAGCACTGGAATCTGACATCTGC 512
Db 421 GTTGTGCGCATGTGCGGAGAAATCTCACAAGCAGCACTGGAATCTGACATCTGC 480
QY 513 CAGCTGTGCGCGCTCTGTGACCCAGTATGAGGCTCGAGAGATGTG 558
Db 481 CAGCTGTGCGCGCTCTGTGACCCAGTATGAGGCTCGAGAGATGTG 526

RESULT 9
US-09-907-372-17/c
; Sequence 17, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Macaca fascicularis
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 703193780J1
US-09-907-372-17

Query Match 24.2%; Score 479.2; DB 10; Length 618;
Best Local Similarity 92.1%; Pred. No. 6.2e-127;
Matches 527; Conservative 0; Mismatches 43; Indels 2; Gaps 2;

QY 1411 ATGACTGAGTCTGAGAAAGCAGAGAGAGGGGCGACAGGACCTTCCCTT 1470
Db 618 ATGACTGAGTCTGAGAAAGCAGAGAGAGGGGCGACAGGACCTTCTT-CTT 560
QY 1471 GAGGCTGCTGCGCCAGCTGGGATTACAGAGGGCTGATGAGGCCCGGGAAGCAGAGC 1530
Db 559 GAGGCTGCTGCGCCAGCTGGGATTACAGAGGGGCTGATGAGGCCCGGGAAGCAGAGC 500
QY 1531 CCTAAGGATTAAGGCTCAGACACTCTGAGAGCAGGTGGGCTGAGTACGGTGC 1590
Db 499 CCTAAGGATTAAGGCTCAGACACTCTGAGAGCAGGTGGGCTGAGTACGGTGC 440

QY 1591 CCTCCACAGACTCTCCCTACTGCTGACGAAACCTGAGGCTCCCGGAGACCCACCCA 1650
Db 439 CCTCCACAGACTCTCCCTACTGCTGACGAAACCTGAGGCTCCCGGAGACCCACCCA 380
QY 1651 CCCCCGAGGCTGCTCAGCTCAGGAGCAGGAGCAGCATGATACCACTGCTGCCAC 1710
Db 379 CCCCAGTGGGCTGCTCAGCTCAGGAGCAGGAGCAGCATGATACCACTGCTGCCAC 320
QY 1711 TACAGCAGGCGCAGCCGAGAGCAGGAGCAGCAGGAGGCGCCACAGAGGCTGCTGAGGA 1770
Db 319 TATGAGCAGGCGCAGCCGAGAGCAGGAGCAGCAGGAGGCGCCACATGCTGCTGAGGA 260
QY 1771 CGTACGAGGCGCTCTTAAGAGATTGCTGCTCATCCCAAGCTTCAGAGACCTTTGG 1830
Db 259 CGTACGAGGCGCTCTTAAGAGATTGCTGCTCATCCCAAGCTTCAGAGACCTTTGG 200
QY 1831 GCTTCCACTTTCAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1890
Db 199 GCTTCCACTTTCAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 140
QY 1891 CGCTCCTTCCCTCCCTCTCTAAGAGAGAGAGAAAGAGTCAATTAACACTAGGGGTTGG 1950
Db 139 CGCTCCTTCCCTCCCTCTCTAAGAGAGAGAGAAAGAGTCAATTAACACTA-GGGTTGG 81
QY 1951 GTAGGATTCCTAGGTATGAGGAGAGATTGG 1982
Db 80 GTAGGATTCCTAGGTATGAGGAGAGATTGG 49

RESULT 10
US-09-919-580-10
; Sequence 10, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7, 40, 356, 492, 507
; OTHER INFORMATION: n = A,T,C or G
US-09-919-580-10

Query Match 22.4%; Score 444.6; DB 10; Length 527;
Best Local Similarity 88.1%; Pred. No. 4.4e-117;
Matches 518; Conservative 0; Mismatches 9; Indels 61; Gaps 1;

QY 47 GAGAGCCCTGAGAGCCGCGGCTGCGGCTCCGCGCTCGGGGTGAGCATGCGCTGTG 106
Db 1 GGGAGCNCCTGAGAGCCGCGGCTGCGGCTCCGCGCTCGNGTGCACATCGGCTGTGAGT 60
QY 107 CCGTCCAGAGCTCTGAGGCTGCGGAGCGCGCCAGCGCTGCGCAGAGCTGCGGCGCTC 166
Db 61 CCGTCCAGAGCTCTGAGGCTGCGGAGCGCC----- 89
QY 167 CTGCTTCTCTCCAGGCCCCACGCTGTGCGCGCTGCGGAGTGGCGCATGCTCT 226
Db 90 -----GCGGCTGCGGAGTGGCGCATGCTCTCT 119
QY 227 GCTTGGGCGACACTGTGCGCGCGGCTGCGGAGCTCTGTGCTGAGGCTCTTGG 286
Db 120 GCTTGGGCGACACTGTGCGCGCGGCTGCGGAGCTCTGTGCTGAGGCTCTTGG 179

Db 1441 AGGGGCCCCAGAGAACCAATTTATACCCATGATGACGAGTCTGTGAGAAAAGCAGAA 1500
Oy 1443 AGGGGGGACAAAGGACCTTCTCCCTTGAGGCTGCTGCCACAGTGGATTCACAGG 1502
Db 1501 AGGGGGGACAAAGGACCTTCTCCCTTGAGGCTGCTGCCACAGTGGATTCACAGG 1560
Oy 1503 GCGTGTAGTAGGGCCCCGGGGAGACAGAGCCCTAAGGGATTTAAGGCTCAGACACCTCTGAGA 1562
Db 1561 GCGTGTAGTAGGGCCCCGGGGAGACAGAGCCCTAAGGGATTTAAGGCTCAGACACCTCTGAGA 1620
Oy 1563 GCAGTGTGGGCACTGCTGGGTAGTACGCTGCTCCACAGACCTCTCCCTACTGCTAGTACA 1622
Db 1621 GAGGTGTGGGCACTGCTGGGTAGTACGCTGCTCCACAGACCTCTCCCTACTGCTAGTACA 1680
Oy 1623 ACCTGAGGCTCTCCGGGACAGACCACCCCTGGGGCTGCTCAGCTCAGGCAAGGAC 1682
Db 1681 ACCTGAGGCTCTCCGGGACAGACCACCCCTGGGGCTGCTCAGCTCAGGCAAGGAC 1740
Oy 1683 AGGGCACTGATACCACTGCTGCTGCCACTACAGACGCGCCACCGGAGCAGGACCGAG 1742
Db 1741 AGGGCACTGATACCACTGCTGCTGCCACTACAGGCGCGCACGCGACAGCGACCGAG 1800
Oy 1743 GGAGCGGCGACACGCTGCTCAGAGAGAGTACGAGGCGCGCTCTAAGGATTCGTGTGC 1802
Db 1801 GGAGCGGCGACACGCTGCTCAGAGAGAGTACGAGGCGCGCTCTAAGGATTCGTGTGC 1860
Oy 1803 TCATCCCAAGCTTTCAGAGACCCCTTTGGGGTTCACACTTCACGTGAGTGAAGTAC 1862
Db 1861 TCATCCCAAGCTTTCAGAGACCCCTTTGGGGTTCACACTTCACGTGAGTGAAGTAC 1920
Oy 1863 CTGCACTGAGATGAAATATATAGGAGAGACGCTCTCCCTCCCTCTAGAGAGAGAA 1922
Db 1921 CTGCACTGAGATGAAATATATAGGAGAGACGCTCTCCCTCCCTCTAGAGAGAGAA 1980
Oy 1923 AGGAGTATTTACACTAGGGGGTTGGGTAGATTTCCAGTATGGGGAAGATTTTGG 1982
Db 1981 AGGAGTATTTACACTAGGGGGTTGGGTAGATTTCCAGTATGGGGAAGATTTTGG 2040

RESULT 5

US-09-907-372-12/c
; Sequence 12, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 862
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 857691871
US-09-907-372-12

Query Match 32.5%; Score 645; DB 10; Length 862;
Best Local Similarity 92.4%; Pred. No. 3.8e-174;
Matches 790; Conservative 0; Mismatches 50; Indels 15; Gaps 10;

Oy 1131 GTGGGCCACGGTACCAATGCGATTCATGTCACGGGGGTCTATGACTATTCACCTGGCAC 1190
Db 853 GTGGGCCCGGGTACCCAAAGCAATAGGT-ACCGGGGGTTCAAGAACATTCACGGGCAAA 795
Oy 1191 A--TCTACATCTACATGAGCA--GTACTGGGGGAGCACCGGGTCTGTGAGACTCTCC 1246
Db 794 ACTGCAATATCAACAAGGACCAAGTACTGGGGGAGACACCGGGGTCTGTGAGACTCTCC 735

Oy 1247 AGTACCCCGGACACC--TCCATACCCCTATCCCGAAGGGGAGACCCCTGCCCTCCGGG 1304
Db 724 AGGCTACCCCGGACCCCTTCCATACCCCTATCCCGAAGAGAGGACCCCTGCCCTCCGGG 675
Oy 1305 CTCTTACACCCCGACAGAGAGATGG--CAAGGCTTGGCACCCTAGCGAGACAGAGC-AC 1361
Db 674 CTCTTACACCCCGACAGAGAGATGGGCAAGGCTTGGCACCCTAGCGAGAGAGAGC-AC 615
Oy 1362 TGTGTGTCACACCCCTTACAGAGGGGGCCCAAGAGAACCAATTTATACCATATGACTGTG 1421
Db 614 TGTGTGTCACACCCCTTACAGAGGGGGCCCAAGAGAACCAATTTATACCATATGACTGTG 555
Oy 1422 AGTGTGAGAAAAGGAGAGAGAGGGGGGACAAAGGGACCTTCTCCCTTGAAGCTGCCCT 1481
Db 554 AGTGTGAGAAAAGGAGAGAGAGGGGGGACAAAGGGACCTTCTCCCTTGAAGGCTGCCCT 495
Oy 1482 GCGCCAGTGGGATTTACAGGGGGGCTGAGTATAGGGGGGGGGAAGGACCCCTAAGGGGAT 1541
Db 494 GCGCCAGTGGGATTTACAGGGGGGCTGAGTATAGGGGGGGGGAAGGACCCCTAAGGGGAT 435
Oy 1542 AAGGCTCAGACACCTGTGAGAGAGAGTGGGCACTGGCTGGGTAGCGTCCCTCCACAGA 1601
Db 434 AAGGCTCAGACACCTGTGAGAGAGAGTGGGCACTGGGTGGGTAGCGTCCCTCCACAGA 375
Oy 1602 CTCTCTCTACTGCTCTGAGCAAACTGAGGCTCTCCGGGAGACCCACCCACCCCTGGGGG 1661
Db 374 CTCTCTCTACTGCTCTGAGCAAACTGAGGCTCTCCGGGAGACCCACCCACCCCTGGGGG 316
Oy 1662 TGCTAGGCTCAGGACGAGGACAGGCGACATGATATCAACTGCTGCTCCACTACAGCAGCC 1721
Db 315 TGCTAGGCTCAGGACGAGGACAGGCGACATGATATCAACTGCTGCTCCACTACAGCAGCC 256
Oy 1722 GCACCGGAGCAGGCG-ACCAGAGGAGCGCCACACAGGTCACCTGCAAGAGAGTCACGGGG 1780
Db 255 GCACCGGAGCAGGCGACGAGCGAGGAGCGCCACAGGTCACCTGCAAGAGAGTCACGGGG 196
Oy 1781 CCCTCTAAGGATTTGTGTGCTCATTCGCCAAGCTTCAGAGACCTTTGGGGTTCACAG 1840
Db 195 CCCT-CTAAGGATTTGTGTGCTCATTCGCCAAGCTTCAGAGACCTTTGGGGTTCACAG 137
Oy 1841 TTCACGTGAGTGAAGTGAACCTGATGATGAATATATAGGAGAGAGCTCTCTCC 1900
Db 136 TTCACGTGAGTGAAGTGAACCTGATGATGAATATATAGGAGAGAGCTCTCTCC 77
Oy 1901 CTCCCTCTCTAGAGAGAGAGAGAGAGTCAATTAACA--CTAGGGGGTGGGTAGGATT 1958
Db 76 CTCCCTCTCTAGAGAGAGAGAGAGAGTCAATTAACAAGTGGGGTGGGTAGGATT 17
Oy 1959 CCTAGTATGGGAA 1973
Db 16 CCTAGTATAGGGGA 2

RESULT 6

US-09-907-372-7/c
; Sequence 7, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7716364H1

OY	1743	GGAGCGCCGACACGGTCACTGTGCAGAGAGACGTACGGGCCCTTAAGATTCGTGGTGC	1802
Dd	1801	GGAGCGCCGACACGGTCACTGTGCAGAGAGACGTACGGGCCCTTAAGATTCGTGGTGC	1860
OY	1803	TGATCCCCAACGCTTCAGAGACCCTTTGGGGGTTCACACTTCACGTGGAATGAGTAGACC	1862
Dd	1861	TGATCCCCAACGCTTCAGAGACCCTTTGGGGGTTCACACTTCACGTGGAATGAGTAGACC	1920
OY	1863	CTGCATGAAAGATGAATTATATAGGAGAGACGCTCCTTCCTCCCCTTCCTAAGAGAAGAA	1922
Dd	1921	CTGCATGAAAGATGAATTATATAGGAGAGACGCTCCTTCCTCCCCTTCCTAAGAGAAGAA	1980
OY	1923	AGGAGATTAATAACAACACTAGAGGGGCTTGAGTAGATTCCTAGGTATGGGAAAGTTTTGG	1982
Dd	1981	AGGAGATTAATAACAACACTAGAGGGGCTTGAGTAGATTCCTAGGTATGGGAAAGTTTTGG	2040
 RESULT 4 US-09-880-107-2135 ; Sequence 2135, Application US/09880107 ; Patent No. US20020142981A1 ; GENERAL INFORMATION: ; APPLICANT: Horne, Darci T. ; APPLICANT: Vockley, Joseph G. ; APPLICANT: Scherf, Uwe ; APPLICANT: Gene Logic, Inc. ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer ; FILE REFERENCE: 44921-5028-WO ; CURRENT FILING DATE: US/09/880,107 ; PRIOR APPLICATION NUMBER: 2001-06-14 ; PRIOR FILING DATE: 2000-06-14 ; PRIOR APPLICATION NUMBER: US 60/237, 054 ; PRIOR FILING DATE: 2000-10-02 ; NUMBER OF SEQ ID NOS: 3950 ; SOFTWARE: Patentin Ver. 2.1 ; SEQ ID NO 2135 ; LENGTH: 2136 ; TYPE: DNA ; ORGANISM: Homo sapiens ; FEATURE: ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L04270 US-09-880-107-2135			
 Query Match 91.3%, Score 1809.2; DB 10; Length 2136; Best Local Similarity 94.6%; Pred. No. 0; Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1.			
OY	51	GCCCTGGAGAGGCCGCGGCTGCGCCGCTCCGCGGCTGGGGCGCACATCGCCCTGAGAGCCGG	110
Dd	1	GCCCTGGAGAGGCCGCGGCTGCGCCGCTCCGCGGCTGGGGCGCACATCGCCCTGAGAGCCGG	60
OY	111	TCCCAAGGCTGTGGGCTGGGGGACCGCGCCACCGCTGTGCCAGAGAGCTGGGGCTTCCTGC	170
Dd	61	TCCCAAGGCTGTGGGCTGGGGGACCGCGCCACCGCTGTGCCAGAGAGCTGGGGCTTCCTGC	120
OY	171	CTTCTCTCCAGGCCGCCACGTTGCTGAGCCGCTTGCGCGAGTGGCGGCATGCTCTGCT	230
Dd	121	CTTCTCTCCAGGCCGCCACGTTGCTGAGCCGCTTGCGCGAGTGGCGGCATGCTCTGCT	180
OY	231	TGGGCGACCTCTGGCGCCCGGCGCTGGCGTGGGGGACCTGTGCTGTGGGGCTCTTGGGGCT	290
Dd	181	TGGGCGACCTCTGGCGCCCGGCGCTGGCGTGGGGGACCTGTGCTGTGGGGCTCTTGGGGCT	240
OY	291	CTGGCAGCATCGACAGCCCGCAGCGGTGCCTCCATATGCTGCGAGAACAGACTGACAG	350
Dd	241	CTGGCAGCATCGACAGCCCGCAGCGGTGCCTCCATATGCTGCGAGAACAGACTGACAG	300
OY	351	GACCAAGAAAAGAAATACTATATAGAGCCCGACGACACCGCATCTGCTCCCGCTGCCGCA	410
Dd	301	GACCAAGAAAAGAAATACTATATAGAGCCCGACGACACCGCATCTGCTCCCGCTGCCGCA	360
OY	411	GGCACCTATGCTCAGCTAATATGTAGCCGCAATCCGGGACACAGTTTGTGCCACATGTGCC	470

US-09-962-436-262
; Sequence 262, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689230-75
; CURRENT APPLICATION NUMBER: US/09-962,436
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 262
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-262

Query Match 91.3%; Score 1809.2; DB 10; Length 2136;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

Oy 51 GCCCTGGAGGCGCGGCGCTCCGCGGCTGGGGTGCAACATGGCCCTGAGTCCG 110
Db 1 GCCCTGGAGGCGCGGCGCTCCGCGGCTGGGGTGCAACATGGCCCTGAGTCCG 60
Oy 111 TCCAGAGCTTGGGCTCGGGGAGCCGCGCCACGCTGCCAGAGAGCTGGGGCTCTGC 170
Db 61 TCCAGAGCTTGGGCTCGGGGAGCCGCGCCACGCTGCCAGAGAGCTGGGGCTCTGC 120
Oy 171 CTCTCTCCAGAGCGCCGCGCTGGGCGCTGGGCGCTGGGCGCTCTCGGGCTC 230
Db 121 CTCTCTCCAGAGCGCCGCGCTGGGCGCTGGGCGCTGGGCGCTCTCGGGCTC 180
Oy 231 TGGGCGACCTCGCCCGCGCTGGGCGCTGGGCGCTGGGCGCTCTCGGGCTC 290
Db 181 TGGGCGACCTCGCCCGCGCTGGGCGCTGGGCGCTGGGCGCTCTCGGGCTC 240
Oy 291 CTGGGAGCATGGAGCCCGCGGCGCTGGGCGCTGGGCGCTGGGCGCTCTCGGGCT 350
Db 241 CTGGGAGCATGGAGCCCGCGGCGCTGGGCGCTGGGCGCTGGGCGCTCTCGGGCT 300
Oy 351 GACCGAGAAAGGAATGATGAGCCCGAGACGCGCATCTGCTCCCGCGCGCGCA 410
Db 301 GACCGAGAAAGGAATGATGAGCCCGAGACGCGCATCTGCTCCCGCGCGCGCA 360
Oy 411 GGCACCTATGCTCAGTAATGTAGCCGATCCGGGACACAGTTTGTGCCATGTC 470
Db 361 GGCACCTATGCTCAGTAATGTAGCCGATCCGGGACACAGTTTGTGCCATGTC 420
Oy 471 GAGAATTCCTACACGAGCACTGGAATGATGATGATGATGATGATGATGATGAT 530
Db 421 GAGAATTCCTACACGAGCACTGGAATGATGATGATGATGATGATGATGATGAT 480
Oy 531 GACCGAGATGAGGCGCTGGAGAGATGCGCCCGAGACGAGAAAGGAGGAGGAG 590
Db 481 GACCGAGATGAGGCGCTGGAGAGATGCGCCCGAGACGAGAAAGGAGGAGGAG 540
Oy 591 GCGTCCAGCGCGGAGATGCTGTGCTGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 650
Db 541 GCGTCCAGCGCGGAGATGCTGTGCTGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 600
Oy 651 CTTTGTGACTGCGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGG 710
Db 601 CTTTGTGACTGCGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGG 660
Oy 711 AACCACTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 770
Db 661 AACCACTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 720

Oy 771 TGCCAGCCCAACACAGGTGTGAGAACCAAGTGTGGAGGAGGAGGCTCCAGGACTGCC 830
Db 721 TGCCAGCCCAACACAGGTGTGAGAACCAAGTGTGGAGGAGGAGGCTCCAGGACTGCC 780
Oy 831 CAGTCCGACACACCTGCAAAAAATTCATTTAGAGCCACTGCCCCAGAGATGTCA----- 884
Db 781 CAGTCCGACACACCTGCAAAAAATTCATTTAGAGCCACTGCCCCAGAGATGTCAAAAC 840
Oy 885 ----- 884
Db 841 ATGCTGATGCTGGCGGTTCTGCTGCACTGCGCTTTTCTGCTCTCCACGCTCTTC 900
Oy 885 -----GATCGCTGCTCAAGAG 902
Db 901 TCCGTCATCTGGAAAGAGCCACCCCTTCTCTGCAAGAAACCTGGATGCTCTCTCAAGAG 960
Oy 903 GGTCCGAGGAGAGAGGAGCCCAATCTGTAGCTGGAGCTGGAGCCCTCCGAGAGCCCAT 962
Db 961 GGTCCGAGGAGAGAGGAGCCCAATCTGTAGCTGGAGCTGGAGCCCTCCGAGAGCCCAT 1020
Oy 963 CCATACCTTCCCTGACTTGTGTACAGGCACTGCTACCCATTCTGGAGATGTTTCCCACTA 1022
Db 1021 CCATACCTTCCCTGACTTGTGTACAGGCACTGCTACCCATTCTGGAGATGTTTCCCACTA 1080
Oy 1023 TCCACTGGGCTCCCGCGAGCCCAATTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1082
Db 1081 TCCACTGGGCTCCCGCGAGCCCAATTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
Oy 1083 CTGACCTGACAGGAGGAGCCGAGTTGGAGACCCGAGGAGAGAGCCAGTGGGCCACG 1142
Db 1141 CTGACCTGACAGGAGGAGCCGAGTTGGAGACCCGAGGAGAGAGCCAGTGGGCCACG 1200
Oy 1143 ACCAATGGCATTCATGATACCGGCGGCTGATGATATACATGCAATGATATACATGAT 1202
Db 1201 ACCAATGGCATTCATGATACCGGCGGCTGATGATATACATGCAATGATATACATGAT 1260
Oy 1203 AATGAGCAGTACTGGGGGAGACACCGGGCTCTGGAGACCTCCAGTACCCCGGAACT 1262
Db 1261 AATGAGCAGTACTGGGGGAGACACCGGGCTCTGGAGACCTCCAGTACCCCGGAACT 1320
Oy 1263 CCATACCCCATTTCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1322
Db 1321 CCATACCCCATTTCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
Oy 1323 GAAGTGGCAAGGCTTGGCACCTAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1382
Db 1381 GAAGTGGCAAGGCTTGGCACCTAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Oy 1383 AGGGGCCCAAGGAACCAATTTATCACCCATGATGATGATGATGATGATGATGATGAT 1442
Db 1441 AGGGGCCCAAGGAACCAATTTATCACCCATGATGATGATGATGATGATGATGATGAT 1500
Oy 1443 AGGGGGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1502
Db 1501 AGGGGGGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Oy 1503 GCGTGAATAGGGCGCGGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1562
Db 1561 GCGTGAATAGGGCGCGGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
Oy 1563 GCAGGTGGGCACTGGCTGGGTACGGTGCCTCCACAGGAGCTTCCCTACTGCTGAGCAA 1622
Db 1621 GCAGGTGGGCACTGGCTGGGTACGGTGCCTCCACAGGAGCTTCCCTACTGCTGAGCAA 1680
Oy 1623 ACCTAGAGGCTTCCCGGAGAGCCACCCAGCCCTGGGGCTGCTGAGGAGGAGGAGGAG 1682
Db 1681 ACCTAGAGGCTTCCCGGAGAGCCACCCAGCCCTGGGGCTGCTGAGGAGGAGGAGGAG 1740
Oy 1683 AGGGCAGATGATACCAATGCTGCTCCACTAGAGAGAGAGGAGGAGGAGGAGGAGGAG 1742
Db 1741 AGGGCAGATGATACCAATGCTGCTCCACTAGAGAGAGAGGAGGAGGAGGAGGAGGAG 1800

QY 241 CTGCCCCGGGCTGGCTGGGGGCTGTGTGCTGGGCTCTTGGGGCTCTGGGAGCAT 300
 DB 241 CTGCCCCGGGCTGGCTGGGGGCTGTGTGCTGGGCTCTTGGGGCTCTGGGAGCAT 300
 QY 301 CGGACCCCCGAGGCTGGCTCTATATGCTGGGAGAACGACCTGGCAGGAGCAGAAA 360
 DB 301 CGGACCCCCGAGGCTGGCTCTATATGCTGGGAGAACGACCTGGCAGGAGCAGAAA 360
 QY 361 AGGAATATATGAGCCGACGACGACATCTGTCTCCGCGGCGGCGGACGACCTATNG 420
 DB 361 AGGAATATATGAGCCGACGACGACATCTGTCTCCGCGGCGGCGGACGACCTATNG 420
 QY 421 TCTCAGTAAATGTAGCCGATCCGGGACACAGTTTGTGCCACATGTCCGAGAAATTCCT 480
 DB 421 TCTCAGTAAATGTAGCCGATCCGGGACACAGTTTGTGCCACATGTCCGAGAAATTCCT 480
 QY 481 ACAAGAGACATGGAATACCTGACCAATCTGCGACGCTGTGCCGCTGTGACCCAGTGA 540
 DB 481 ACAAGAGACATGGAATACCTGACCAATCTGCGACGCTGTGCCGCTGTGACCCAGTGA 540
 QY 541 TGGGCTGGAGGAGATTCGCCCCCTGCACAGAAACGAGAACCGAGTGGCGCTGCCAGC 600
 DB 541 TGGGCTGGAGGAGATTCGCCCCCTGCACAGAAACGAGAACCGAGTGGCGCTGCCAGC 600
 QY 601 CGGGAATGTCTGTGCTGCTGCTGGGCTCTGAGTGTACACACTGCGAGTACTTCTGACT 660
 DB 601 CGGGAATGTCTGTGCTGCTGCTGGGCTCTGAGTGTACACACTGCGAGTACTTCTGACT 660
 QY 661 GCGGCTGGGACCTGAGAGCCGAGCTCAAGATGAAGTTGGAGGGTAAACAACACTGGC 720
 DB 661 GCGGCTGGGACCTGAGAGCCGAGCTCAAGATGAAGTTGGAGGGTAAACAACACTGGC 720
 QY 721 TCCCTGGAAGGCGAGGAGCACTTCAGAAATACCTCTCCCGGCGGCTGGCCAGCCCC 780
 DB 721 TCCCTGGAAGGCGAGGAGCACTTCAGAAATACCTCTCCCGGCGGCTGGCCAGCCCC 780
 QY 781 ACACCAAGGTGTGAGAACCAAGGTGTGTGAGAGAGCTCCAGGCACTGCCAGTCCGACA 840
 DB 781 ACACCAAGGTGTGAGAACCAAGGTGTGTGAGAGAGCTCCAGGCACTGCCAGTCCGACA 840
 QY 841 CAACCTGCAAAAATCCATTAGAGCACTGCCCGACAGAGATGTGAGAGTCTGCTCAAGA 900
 DB 841 CAACCTGCAAAAATCCATTAGAGCACTGCCCGACAGAGATGTGAGAGTCTGCTCAAGA 900
 QY 901 GGGCTGCCGAGGAGAGGAGACCCATCTGTAGCTGGAAGCTGGAGGAGCTCCGAGGGCC 960
 DB 901 GGGCTGCCGAGGAGAGGAGACCCATCTGTAGCTGGAAGCTGGAGGAGCTCCGAGGGCC 960
 QY 961 ATCATATCTTCCCTGACTGTGTAGACCACTGCTACCCATTTCTGAGATGTTTCCCGAG 1020
 DB 961 ATCATATCTTCCCTGACTGTGTGTAGACCACTGCTACCCATTTCTGAGATGTTTCCCGAG 1020
 QY 1021 TATCCACTGGGCTCCCGGACGCGCAGTTTGGAGGAGGGGTGCCGAACGACAGATC 1080
 DB 1021 TATCCACTGGGCTCCCGGACGCGCAGTTTGGAGGAGGGGTGCCGAACGACAGATC 1080
 QY 1081 CTTCTGAGCTGTGACGAGGAGCGGAGTTGGAAACCCGGGAGAGAGCCAGGGTGGCCAG 1140
 DB 1081 CTTCTGAGCTGTGACGAGGAGCGGAGTTGGAAACCCGGGAGAGAGCCAGGGTGGCCAG 1140
 QY 1141 GTACCAATGGCAATTCATGTACACGGGCGGTATGACTATCATCTGGAACATCTACATCT 1200
 DB 1141 GTACCAATGGCAATTCATGTACACGGGCGGTATGACTATCATCTGGAACATCTACATCT 1200
 QY 1201 ACAATGAGACATGAGTGGGGGAGACACGGGCTCTGAGAGCTCCAGGTACCCCGAAC 1260
 DB 1201 ACAATGAGACATGAGTGGGGGAGACACGGGCTCTGAGAGCTCCAGGTACCCCGAAC 1260
 QY 1261 CTCATATACCCCTATTCGCGAAGAGGGGAGACCTGGGCTCCGGGGCTCTTACACCCAGC 1320
 DB 1261 CTCATATACCCCTATTCGCGAAGAGGGGAGACCTGGGCTCCGGGGCTCTTACACCCAGC 1320
 QY 1321 AGGAAGATGGCAAGGCTTGGCACTAGCGGAGACAGACATGTGTGTCACACCTCTTA 1380

DB 1321 AGGAAGATGGCAAGGCTTGGCACTAGCGGAGAGACACTGTGTGCTCCACACCTCTA 1380
 QY 1381 ACAGGGCCCCAAGAGCAACATTTATACACCATGACTGACTGTGAGAAAAGGCAGAA 1440
 DB 1381 ACAGGGCCCCAAGAGCAACATTTATACACCATGACTGACTGTGAGAAAAGGCAGAA 1440
 QY 1441 GAAGGGGGGCAAGGGACCTTCTCCCTTGAAGGTGCTGCTGCGCCAGCTGGGATTCACAG 1500
 DB 1441 GAAGGGGGGCAAGGGACCTTCTCCCTTGAAGGTGCTGCTGCGCCAGCTGGGATTCACAG 1500
 QY 1501 GGGCTGAGTATGAGGCTCCGGGAGAGAGACCCCTAAGGATTAAGGCTCAGACCTCTGA 1560
 DB 1501 GGGCTGAGTATGAGGCTCCGGGAGAGAGACCCCTAAGGATTAAGGCTCAGACCTCTGA 1560
 QY 1561 GAGCAGGTGGGACACTGCTGGGTACGAGTGGCTCCACAGAGACTCTCCCTACTGCTGAGC 1620
 DB 1561 GAGCAGGTGGGACACTGCTGGGTACGAGTGGCTCCACAGAGACTCTCCCTACTGCTGAGC 1620
 QY 1621 AAACCTGAGGCTCTCCGGGACAGCCACACCCCTGGGGGCTGCTCAGCTCAGGACAGG 1680
 DB 1621 AAACCTGAGGCTCTCCGGGACAGCCACACCCCTGGGGGCTGCTCAGCTCAGGACAGG 1680
 QY 1681 ACAGGGACATGATACCACTGCTGCCACTACAGACAGCCGACCGAGGACAGCAGCAGC 1740
 DB 1681 ACAGGGACATGATACCACTGCTGCCACTACAGACAGCCGACCGAGGACAGCAGCAGC 1740
 QY 1741 AGGAGCCGCCACACGCTACCTGCAAGAGAGCTCAAGGAGGCTTAAAGATTTGTGTGT 1800
 DB 1741 AGGAGCCGCCACACGCTACCTGCAAGAGAGCTCAAGGAGGCTTAAAGATTTGTGTGT 1800
 QY 1801 GCTCATCCCAAGCTTCAAGAGACCTTGGGGTTCACACTACAGTGTGAGTGTGAGTGA 1860
 DB 1801 GCTCATCCCAAGCTTCAAGAGACCTTGGGGTTCACACTACAGTGTGAGTGTGAGTGA 1860
 QY 1861 CCCTCATGAGAGTAAATTTATAGGAGAGACCTCTTCCCTCCCTCTAGAGAGAGAG 1920
 DB 1861 CCCTCATGAGAGTAAATTTATAGGAGAGACCTCTTCCCTCCCTCTAGAGAGAGAG 1920
 QY 1921 AAAGGAGCTATTAAACAACAGAGGGGTTGGGTAGGATTCATGATGTGGGAGAGCTTTT 1980
 DB 1921 AAAGGAGCTATTAAACAACAGAGGGGTTGGGTAGGATTCATGATGTGGGAGAGCTTTT 1980
 QY 1981 GG 1982
 DB 1981 GG 1982

RESULT 2
 US-09-954-531-1348
 ; Sequence 1348, Application US/09954531
 ; Patent No. US20020165180A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weaver, Zoe
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
 ; FILE REFERENCE: 689290-77
 ; CURRENT FILING DATE: 2002-05-02
 ; PRIOR APPLICATION NUMBER: US/09/954,531
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,133
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,009
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,034
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,509
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,567
 ; NUMBER OF SEQ ID NOS: 1392
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1348
 ; LENGTH: 2136

[illegible]

RC TISSUE-Lung cancer;
 RX MEDLINE=98151033; PubMed=9492069;
 RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
 RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
 RA Tsuda E., Morinaga T., Higashio K.;
 RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and
 RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
 RT osteoclastogenesis in vitro.";
 RL Endocrinology 139:1329-1337(1998).
 [3]
 RP SEQUENCE FROM N.A., AND VARIANT ASN-3.
 RC TISSUE-Placenta;
 RX MEDLINE=9831569; PubMed=9688283;
 RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
 RT "Cloning and characterization of the gene encoding human
 RT osteoprotegerin/osteoclastogenesis-inhibitory factor.";
 RL Eur. J. Biochem. 254:685-691(1998).
 [4]
 RP SEQUENCE FROM N.A., AND VARIANT ASN-3.
 RC TISSUE-Eye;
 RX Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 [5]
 RP SEQUENCE OF 22-36 AND 378-401.
 RX MEDLINE=98238645; PubMed=9571159;
 RA Tomoyasu A., Goto M., Fujise N., Mochizuki S.-I., Yasuda H.,
 RA Morinaga T., Tsuda E., Higashio K.;
 RT "Characterization of monomeric and homodimeric forms of
 RT osteoclastogenesis inhibitory factor.";
 RL Biochem. Biophys. Res. Commun. 245:382-387(1998).
 [6]
 RP SEQUENCE OF 22-393 FROM N.A.
 RC TISSUE-Placenta;
 RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
 RT "Cloning and expression of osteoprotegerin from Homo sapiens.";
 RL Acta Biochim. Biophys. Sin. 31:680-684(1999).
 [7]
 RP SEQUENCE OF 242-255; 354-359 AND 369-378, AND FUNCTION.
 RX MEDLINE=97312536; PubMed=9168977;
 RA Tsuda E., Goto M., Mochizuki S.-I., Yano K., Kobayashi F.,
 RA Morinaga T., Higashio K.;
 RT "Isolation of a novel cytokine from human fibroblasts that
 RT specifically inhibits osteoclastogenesis.";
 RL Biochem. Biophys. Res. Commun. 234:137-142(1997).
 [8]
 RP TRAIL BINDING.
 RX MEDLINE=98269100; PubMed=9603945;
 RA Emery J.G., McDonnell P., Burke M.B., Deen K.C., Lyn S., Silverman C.,
 RA Dul E., Appelbaum E.R., Eichman C., DiPrinzio R., Dadds R.A.,
 RA James I.E., Rosenberg M., Lee J.C., Young P.R.;
 RT "Osteoprotegerin is a receptor for the cytotoxic ligand TRAIL.";
 RL J. Biol. Chem. 273:14363-14367(1998).
 [9]
 RP CHARACTERIZATION, AND MUTAGENESIS OF CYS-400.
 RX MEDLINE=98148058; PubMed=9478964;
 RA Yamaguchi K., Kinosaki M., Goto M., Kobayashi F., Tsuda E.,
 RA Morinaga T., Higashio K.;
 RT "Characterization of structural domains of human osteoclastogenesis
 RT inhibitory factor.";
 RL J. Biol. Chem. 273:5117-5123(1998).
 [10]
 RP REVIEW.
 RX MEDLINE=21395914; PubMed=11505389;
 RA Hofbauer L.C., Neubauer A., Heufelder A.E.;
 RT "Receptor activator of nuclear factor-kappaB ligand and
 RT osteoprotegerin: potential implications for the pathogenesis and
 RT treatment of malignant bone diseases.";
 RL Cancer 92:460-470(2001).
 -1- FUNCTION: Acts as decoy receptor for RANKL and thereby neutralizes
 its function in osteoclastogenesis. Inhibits the activation of
 osteoclasts and promotes osteoclast apoptosis in vitro. Bone
 homeostasis seems to depend on the local RANKL/OPG ratio. May also
 play a role in preventing arterial calcification. May act as decoy

receptor for TRAIL and protect against apoptosis. TRAIL binding
 blocks the inhibition of osteoclastogenesis.
 -1- SUBUNIT: Homodimer.
 -1- SUBCELLULAR LOCATION: Secreted.
 -1- TISSUE SPECIFICITY: Highly expressed in adult lung, heart, kidney,
 liver, spleen, thymus, prostate, ovary, small intestine, thyroid,
 lymph node, trachea, adrenal gland, testis, and bone marrow.
 Detected at very low levels in brain, placenta and skeletal
 muscle. Highly expressed in fetal kidney, liver and lung.
 -1- INDUCTION: Upregulated by increasing calcium concentration in the
 medium and estrogens. Downregulated by glucocorticoids.
 -1- PTM: N-glycosylated. Contains sialic acid residues.
 -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 -1- SIMILARITY: CONTAINS 2 DEATH DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: U94332; AAB53709.1; -
 CC EMBL: AB002146; BAA25910.1; -
 CC EMBL: AB008822; BAA32076.1; -
 CC EMBL: AB008821; BAA32076.1; JOINED.
 CC EMBL: BC030155; AAH30155.1; -
 CC EMBL: AF134187; AAF20168.1; -
 CC HSSP: P25942; ICDF.
 CC GeneW: HGNC:11909; TNFRSF11B.
 CC MIM: 602643; -
 DR InterPro: IPR000488; Death.
 DR InterPro: IPR001368; TNFR_c6.
 DR Pfam: PF00020; TNFR_c6; 3.
 DR ProDom: PD000771; TNFR_c6; 1.
 DR SMART: SM00208; TNFR; 4.
 DR SMART: SM00205; DEATH; 1.
 DR PROSITE: PS50017; DEATH DOMAIN; FALSE_NEG.
 DR PROSITE: PS00652; TNFR_NGFR_1; 2.
 DR PROSITE: PS50050; TNFR_NGFR_2; 2.
 KW Receptor; Apoptosis; Glycoprotein; Repeat; Signal; Polymorphism.
 FT SIGNAL 1 21
 FT CHAIN 22 401
 FT
 FT REPEAT 24 62
 FT REPEAT 65 105
 FT REPEAT 107 142
 FT REPEAT 145 185
 FT DOMAIN 198 269
 FT DOMAIN 270 365
 FT SITE 400 400
 FT DISULFID 41 54
 FT DISULFID 44 62
 FT DISULFID 65 80
 FT DISULFID 83 97
 FT DISULFID 87 105
 FT DISULFID 107 118
 FT DISULFID 124 142
 FT DISULFID 145 160
 FT DISULFID 145 160
 FT DISULFID 166 185
 FT CARBOHYD 98 98
 FT CARBOHYD 152 152
 FT CARBOHYD 165 165
 FT CARBOHYD 178 178
 FT CARBOHYD 289 289
 FT VARIANT 3 3
 FT
 FT MUTAGEN 400 400
 FT MUTAGEN 401 401
 FT CONFLICT 263 263
 FT SEQUENCE 401 AA; 46040 MW; EDF448B67D86C71E CMC64;

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RA Isogai D., Ichino M., Yoshinari M., Yamaura A., Kurokawa F.,
 RA Minami M.;
 RT "Mouse DR6: mouse homolog of human TNFR-related death receptor-6
 (DR6).";
 RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Kidney;
 RA Kim V., Machleidt T., Shi W.-X., Wang X., Cai Z.;
 RT "Murine DR6: murine TNFR-related death receptor-6";
 RT Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strusberg R.;
 RT Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP FUNCTION.
 RX MEDLINE=21571606; PubMed=11714751;
 RA Zhao H., Yan M., Wang H., Erickson S., Grewal I.S., Dixit V.M.;
 RT "Impaired c-Jun amino terminal kinase activity and T cell
 RT differentiation in death receptor 6-deficient mice.";
 RL J. Exp. Med. 194:1441-1448(2001).
 CC -1- FUNCTION: May activate NF-kappa-B and promote apoptosis (By
 CC similarity). May activate JNK and be involved in T-cell
 CC differentiation.
 CC -1- FUNCTION: May activate NF-kappa-B and promote apoptosis.
 CC May be involved in T-cell differentiation.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC -1- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in adult spleen,
 CC thymus, testis, prostate, ovary, small intestine, colon, brain,
 CC lung and kidney, and in fetal brain, liver and lung. Detected at
 CC lower levels in adult peripheral blood leukocytes, lung, and in
 CC fetal muscle, heart, kidney, small intestine and skin. Detected in
 CC T-cells, B-cells and monocytes. In T-cells expression is highest
 CC in Th0 cells, intermediate in Th2 cells and lower in Th1 cells.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-25 IS THE INITIATOR.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC -----
 CC EMBL: AF322069; AAC38115.1; -;
 DR EMBL: AY043489; AAK74193.1; -;
 DR EMBL: BC016420; AAH16420.1; -;
 DR HSP: 014763; 1D0G.
 DR MGI: 2151075; Tnfrsf21.
 DR InterPro: IPR000488; Death.
 DR InterPro: IPR001368; TNFR_C6.
 DR Pfam: PF00020; TNFR_C6; 4.
 DR ProDom: PD000771; TNFR_C6; 1.
 DR SMART: SMO0005; DEATH; 1.
 DR SMART: SMO0208; TNFR; 4.
 DR PROSITE: PS50017; DEATH DOMAIN; 1.
 DR PROSITE: PS50052; TNFR_NGFR_1; 1.
 DR PROSITE: PS50050; TNFR_NGFR_2; 1.
 KW Receptor; Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 41
 FT CHAIN 42 655
 FT DOMAIN 42 349
 FT TRANSMEM 350 370
 FT DOMAIN 371 655
 FT DOMAIN 415 498
 FT REPEAT 50 88
 FT TNFR-CYS 1.

FT REPEAT 90 131
 FT REPEAT 133 167
 FT REPEAT 170 211
 FT DISULFID 67 80
 FT DISULFID 70 88
 FT DISULFID 91 106
 FT DISULFID 109 123
 FT DISULFID 113 131
 FT DISULFID 133 144
 FT DISULFID 150 168
 FT DISULFID 171 186
 FT DISULFID 192 211
 FT CARBOHYD 82 82
 FT CARBOHYD 141 141
 FT CARBOHYD 252 252
 FT CARBOHYD 257 257
 FT CARBOHYD 278 278
 FT CARBOHYD 289 289
 FT CONFLICT 352 352
 FT CONFLICT 523 523
 FT SEQUENCE 655 AA; 71982 MW; 5EC7C51C7C99EFF7 CRC64;
 SO
 Query Match 10.8%; Score 243; DB 1; Length 655;
 Best Local Similarity 26.5%; Pred. No. 2.9e-07;
 Matches 76; Conservative 37; Mismatches 124; Indels 50; Gaps 13;
 QY 57 ICSSRCPPGTYYSAKSCSRIDYCATCAENSINVEHNNYLTICOLC-RPCD-PWGLCEETIA 114
 DB 65 LNCDCPKACTYVSEHCTNMSLRVCSGCPAGTFREHNGIERCHDCQPCPPWPMI---ERI 121
 QY 115 PCRSKRKTCROCPGME-----CAAWALRCTHCELLSDCPGTEALTKDEVGKGN--HC 167
 DB 122 PCAALTDRECICTPPGMYGNGTCAPHTV-----CPVGVGKK---GTENDVVC 168
 QY 168 VPCKAGHPONTSSPSARCPHRCENQGLVEAPGTAOSDTCTCKNPL--EPLPEMSSGL 225
 DB 169 KCCARGTFFDVSSVYKCKAHTCLDGLQNEVVPKRETDNCGMRLFFSSTNPSSGTV 228
 QY 226 LRRPQG-EGPNVAGSWMPRAHPFPPDLYOPLPISGDVSYSGTLAAPPLVLEAGVQ 284
 DB 229 TRSHPEHMSHDVPSSTYEPQGN-----TDSNSTSVKRVPSG-IEETVPD 277
 QY 285 QOSPLD-----LTREPLEPEQSQVAHGNTGITHVGVGSMITING 323
 DB 278 NNSSTGKGTCTNLTLPNPQV---THQAPHHHILKLKLPSSMELATG 321
 RESULT 14
 TR14_HUMAN
 AC 092956; Q9UM65; Q96J31; Q8WKRL;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 14 precursor
 DE (Herpesvirus entry mediator A) (Tumor necrosis factor receptor-like 2)
 DE (TR2).
 GN TNPFRS14 OR HVEM OR HVEA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cervical adenocarcinoma;
 RX MEDLINE=97053782; PubMed=8898196;
 RA Montgomery R.L., Warner M.S., Lum B.J., Spear P.G.;
 RT "Herpes simplex virus-1 entry into cells mediated by a novel member of
 RT the TNF/NGF receptor family.";
 RL Cell 87:427-436(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97306336; PubMed=9162061;

Query Match	Best Local Similarity	Score	DB	Length
Matches 105; Conservative	42; Mismatch	152; Indels	117; Gaps	22;
7 TSAPGLANGSVL-----GLFGFLAASQPAVPAASENQTCDQKEYEPQHR	56	10.96	244.5	461
4 STVEPLL-PLVLRALVDVYAGVHGLV-----LHPGDRKRESLCPOGKISHPNR	55	25.28	1.8e-07	
57 -TCSRCPPTGYVSAKC-SRIIDTVCAACAEVSYNEHNNYLTICOLCRPCDPVGLLETA	114			
56 SICCKCKHKGYLIHNDCLGPGIDTCRCRCDNQTFTASENHLLQCLSCSKCRSEMSQVEIS	115			
115 PCTSKRKTQCRQCPGMCFAANA--LECTHCELSDCPPGE---AEIKDEVGKGNHNC	167			
116 PCTVDRDVTGCCRKNQYKRKYWSEFLQCLNCSL--CNGTVQLPCLKEKODTI-----	165			
168 VPCKAGHQTNTSSPARCQPHTRCENOGVLEAPGTAQ-----SDTCKKNPL-----	214			
166 CNCHSGEFLR---DKCEVSCVNCNADCKNLCPATSETRRNDPQTGTVLLPLVIFGEL	221			
215 -----EPLRPENSGSL-RRPDQEG-PNPVAGSWEPKRAHPYFDPLOYPL	258			
222 CLAFPLFVGLACRQGRMKPKLYSIICGSTVYKBESEPELATA---PSRGP-----ITTF	273			
259 LPISDVSPVSTGLPAALVLEAGVQOQSPDLITREPQLREPEQSAVHGTNGIHVTS	318			
274 SPIP-SFSPPTTFSP-----YP-SFSPIS--SPFTFPDWS-----	305			
319 MTITNITIVNGPVLGPPGPDLPATPEP-----PYPIPEGGDPGPGGLSTPHQ	368			
306 -----NIVTSPPKIAPPPQAGPILMPASTVTPPLPRKWSGSAHSAHSAHPAQ	356			

KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal;
 KW Alternative splicing.
 FT SIGNAL 1 19
 FT CHAIN 20 289
 FT DOMAIN 20 193
 FT TRANSMEM 194 215
 FT DOMAIN 216 289
 FT REPEAT 25 60
 FT REPEAT 61 103
 FT REPEAT 104 144
 FT REPEAT 145 187
 FT DISULFID 26 37
 FT DISULFID 38 51
 FT DISULFID 41 59
 FT DISULFID 62 77
 FT DISULFID 83 103
 FT DISULFID 105 119
 FT DISULFID 111 116
 FT DISULFID 125 143
 FT CARBOHYD 153 153
 FT VARSPLIC 166 203
 FT VARSPLIC 204 289
 FT VARSPLIC 187 216
 FT VARSPLIC 216 234
 FT VARSPLIC 235 289
 FT VARSPLIC 216 222
 FT VARSPLIC 223 289
 SQ SEQUENCE 289 AA; 32111 MW; C791CB6D2FEA57AE CRC64;
 Query Match 11.3%; Score 253; DB 1; Length 289;
 Best Local Similarity 25.5%; Pred. No. 3.8e-08;
 Matches 82; Conservative 38; Mismatches 128; Indels 74; Gaps 13;
 QY 8 SAPGLA--WGPLVLGLFGLLAQPOAVPYASENQTRDQDEKEYEYEPQHRICSRCPG 65
 DB 3 SLPLRCALMGCL-----LTAVHLGQCV-----TCSDD-KQYLHGOQ--CDQLCPG 44
 QY 66 TVYSAKCSRIRDTVCATCAENSYNEHMYLITCOLCRPDPMGLLEIAPCTSKRKTQCR 125
 DB 45 SRTLSHCALERTQCHPDSCGFSSAQMNRLEIQRHRCRCPQGLRVKKEGTAESDTVCT 104
 QY 126 CQPGKFCAMALECTHCELLSDCPG-----TEALKDEYVGKGNHCYPCRAKHQNTSS 180
 DB 105 CKEGHCT--SKDCRCAQHPCIPGFGVEMEMATEITTV-----CHCPVGFSSNOS 156
 QY 181 PSARCOPTRCENOGILVEAPGTAOSDTTC--KNPLEPL--PREMGSLK----- 227
 DB 157 LFEKCYPTSCEDKNLEVLQKGTSTQNTVICGLSKRMALLVLPVVMGLITITFGVFLYIK 216
 QY 228 ---RRPGEGPNVAGSWEPKRAHPFPLVQPLPLISGDVSPVSTGLPAAPVLEAGVQ 284
 DB 217 KVKKPKDNEMLPARARQDDPEMEDVPG-----HNTAPVQDTLHGCGPV 262
 QY 285 QOSPLDLTRPQLRPGESQVA 306
 DB 263 TQ-----EDGKESRIS 273
 RESULT 10
 TNR5_BOVIN STANDARD: PRT; 269 AA.
 AC Q28203;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 5 precursor
 DE (CD40L receptor) (B-cell surface antigen CD40) (Fragment).
 GN TNFSF5 OR CD40.

OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97281252; PubMed=9135560;
 RA Hirano A., Brown W.C., Estes D.M.;
 RT "Cloning, expression and biological function of the bovine CD40
 RT homologue: role in B-lymphocyte growth and differentiation in
 RT cattle."
 RL Immunology 90:294-300(1997).
 CC -1- FUNCTION: Receptor for TNFSF5/CD40L.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U57745; AAC48710.1; -.
 CC HSSP: P25942; ICDF.
 DR InterPro: IPR001368; TNFR_C6.
 DR Pfam: PF00020; TNFR_C6; 4.
 DR ProDom: PD000771; TNFR_C6; 1.
 DR SMART: SM00208; TNFR; 4.
 DR PROSITE: PS00652; TNFR_NGFR_1; 1.
 DR PROSITE: PS50050; TNFR_NGFR_2; 1.
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 269
 FT DOMAIN 20 193
 FT TRANSMEM 194 215
 FT DOMAIN 216 269
 FT REPEAT 25 60
 FT REPEAT 61 103
 FT REPEAT 104 144
 FT REPEAT 145 187
 FT DISULFID 26 37
 FT DISULFID 38 51
 FT DISULFID 41 59
 FT DISULFID 62 77
 FT DISULFID 83 103
 FT DISULFID 105 119
 FT DISULFID 111 116
 FT DISULFID 125 143
 FT CARBOHYD 153 153
 FT CARBOHYD 180 180
 FT NON_TER 269
 SQ SEQUENCE 269 AA; 746903F30F95F387 CRC64;
 Query Match 11.2%; Score 251.5; DB 1; Length 269;
 Best Local Similarity 33.1%; Pred. No. 4.3e-08;
 Matches 55; Conservative 21; Mismatches 81; Indels 9; Gaps 4;
 QY 47 EKEYEPQHRICSRCPGVYSAKSRIRDTVCATCAENSYNEHMYLITCOLCRPCDP 106
 DB 28 EKQY--PVNSLCCDCLCPGQKLVNCTEVSSTEQSCGKGFELSTWNEKXCHERYGNP 85
 QY 107 VMGLEIAPCTSKRKTQCRCPGMFCAMALECTHCELLSDCPGTEALKDEYVGK--N 164
 DB 86 NLGLRIQSEGTLNPTTICVVEGQCHSHT--CESCTPHSLCLPQFGVK--QIATGLLD 140
 QY 165 NHCVCKRKGHNQNTSSPSARCOPTRCENOGILVEAPGTAOSDTTC 210
 DB 141 TVCEPCPLGFSNVSSAFKCHRWTSCKRGLVDEHVGNTKTDVVC 186

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL: AF018253; AAB86809.1; -.
 DR HSSP: P25942; ICDF.
 DR Genew: HGNC:11908; TNFRSF11A.
 DR MIM: 603499; -.
 DR MIM: 174810; -.
 DR MIM: 602080; -.
 DR InterPro: IPR001368; TNFR_c6.
 DR Pfam: PF00020; TNFR_c6; 4.
 DR ProDom: PD000771; TNFR_c6; 1.
 DR SMART: SM00208; TNFR; 4.
 DR PROSITE: PS00652; TNFR_NGFR_1; 1.
 DR PROSITE: PS00652; TNFR_NGFR_2; 1.
 DR Receptor: Transmembrane; Glycoprotein; Repeat; Signal; Polymorphism;
 Disease mutation.
 KW SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 616 TUMOR NECROSIS FACTOR RECEPTOR
 FT TRANSMEM 213 233 SUPERFAMILY MEMBER 11A.
 FT DOMAIN 234 616 EXTRACELLULAR (POTENTIAL).
 FT REPEAT 34 68 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 71 112 TNFR-CYS 1.
 FT REPEAT 114 151 TNFR-CYS 2.
 FT REPEAT 154 194 TNFR-CYS 3.
 FT DISULFID 34 46 TNFR-CYS 4.
 FT DISULFID 47 60 BY SIMILARITY.
 FT DISULFID 50 68 BY SIMILARITY.
 FT DISULFID 71 86 BY SIMILARITY.
 FT DISULFID 92 112 BY SIMILARITY.
 FT DISULFID 114 127 BY SIMILARITY.
 FT DISULFID 133 151 BY SIMILARITY.
 FT CARBOHYD 105 105 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC...) (POTENTIAL).
 FT VARIANT 21 21 L -> LALLLCALL (IN PDB2).
 FT VARIANT 21 21 /FtId-VAR_011516.
 FT VARIANT 21 21 L -> LALLLCALL (IN FEO).
 FT VARIANT 21 21 /FtId-VAR_011517.
 FT VARIANT 21 21 A -> V.
 FT VARIANT 192 192 /FtId-VAR_011518.
 FT SEQUENCE 616 AA; 66033 MW; E3DE9A7A08196F81 CRC64;
 SO
 Query Match 13.3%; Score 298; DB 1; Length 616;
 Best Local Similarity 31.1%; Pred. No. 2,1e-10;
 Matches 76; Conservative 31; Mismatches 97; Indels 40; Gaps 7;
 DB 14 LILLICALLARLOVALOIAIPCTSE-----KHYEHLGR-CCKCKEKGKMSKCTTT 63
 QY 18 VLGFLGLAASOP--QAVPPVASENQTRODEKEYEPQHRICGRCRPPGYVSAKCSRI 75
 DB 14 LILLICALLARLOVALOIAIPCTSE-----KHYEHLGR-CCKCKEKGKMSKCTTT 63
 QY 76 RDTVCATCAENSYNHMYVLTICQCRPDPMVGLSEIAPCTSKKRTCCRCQPGMFCAM 135
 DB 64 SDSVCLPCQPDYLDSDWNEDEKCLHKYCDICKALVAVAGSTIPRRACACAGVH--W 120
 QY 136 ALECHTCELLSDCPPTGEATELEKDEYKGNHCVCCKAGHPONTSSPSARCPHTCENOG 195
 DB 121 SODCECCRRNTCAFGIAGOHPLQINK-DTVCKPCIAAGYFSAFSTDKCRFWITCTELG 179
 QY 196 IVEAAPTQSDPTCKKNPLEPLRPMSGSLKRRPQGEPPNVAVASWEPPKAPHPEPDIV 255
 DB 180 KVEHHGTETKSDAVC-----SSSLPRKPP-----PNEHYVLLPGLI 215
 QY 256 QPLL 259
 DB 216 IILL 219
 RESULT 9
 TNRS_MOUSE

 ID TNRS_MOUSE STANDARD: PRT: 289 AA.
 AC P27512; Q99NE0; Q99NE1; Q99NE2; Q99NE3;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 5 precursor
 DE (CD40L receptor) (B-cell surface antigen CD40) (BP50) (CDw40).
 GN TNFRSF5 OR CD40.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP MEDLINE=92105763; PubMed=1370315;
 RX Torres R.M., Clark E.A.,
 RA "Differential increase of an alternatively polyadenylated mRNA
 RT species of murine CD40 upon B lymphocyte activation";
 RL J. Immunol. 148:620-626(1992).
 RN [2]
 RP REVISIONS.
 RC STRAIN-BALB/c;
 RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM I).
 RC STRAIN-BALB/c; TISSUE=Liver;
 RX MEDLINE=93094586; PubMed=1281194;
 RA Girmaldi J.C., Torres R., Kozak C.A., Chang R., Clark E.A.,
 RA Howard M., Cockayne D.A.;
 RT "Genomic structure and chromosomal mapping of the murine CD40 gene";
 RL J. Immunol. 149:3921-3926(1992).
 RN [4]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING (ISOFORMS II; III; IV AND
 V).
 RX MEDLINE=21117110; PubMed=11172023;
 RA Tone M., Tone Y., Fairchild P.J., Wykes M., Waldmann H.;
 RT "Regulation of CD40 function by its isoforms generated through
 RL alternative splicing";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:1751-1756(2001).
 CC - FUNCTION: Receptor for TNFSF5/CD40L.
 CC - SUBCELLULAR LOCATION: Type I membrane protein (isoforms I, III, IV
 and V); Secreted (isoform II).
 CC - ALTERNATIVE PRODUCTS: 5 isoforms; I (shown here), II, III, IV and
 V; are produced by alternative splicing.
 CC - SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

RA Nakagawa N., Kinoshita M., Yamaguchi K., Shima N., Yasuda H., Yano K.,
 RA Morinaga T., Higashio K.;
 RT "RANK is the essential signaling receptor for osteoclast
 differentiation factor in osteoclastogenesis.";
 RL Biochem. Biophys. Res. Commun. 253:395-400(1998).
 CC -1- FUNCTION: Receptor for TNFRSF11/RANKL/TRANCE/OPGL; essential for
 CC RANKL-mediated osteoclastogenesis. Involved in the regulation of
 CC interactions between T-cells and dendritic cells.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS EXPRESSION WITH HIGH LEVELS IN
 CC TRABECULAR BONE, THYMUS, SMALL INTESTINE, LUNG, BRAIN AND KIDNEY.
 CC WEEKLY EXPRESSED IN SPLEEN AND BONE MARROW.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF019046; AAB6810.1; -
 DR EMBL: BC019185; AAB19185.1; -
 DR HSSP: P25942; 1CDF.
 DR MGD: MGI:1314891; Tnf1sfl1a.
 DR InterPro: IPR001368; TNFR.C6.
 DR Pfam: PF000020; TNFR.C6; 3.
 DR Prodom: P0000771; TNFR.C6; 1.
 DR SMART: SM00208; TNFR.4.
 DR PROSITE: PS00652; TNFR_NGFR_1; 1.
 DR PROSITE: PS50050; TNFR_NGFR_2; 1.
 DR Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 30
 FT CHAIN 1 625 TUMOR NECROSIS FACTOR RECEPTOR
 FT DOMAIN 31 214 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 215 234 POTENTIAL.
 FT DOMAIN 235 625 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 35 69 TNFR-CYS 1.
 FT REPEAT 72 113 TNFR-CYS 2.
 FT REPEAT 115 152 TNFR-CYS 3.
 FT REPEAT 155 195 TNFR-CYS 4.
 FT REPEAT 35 47 BY SIMILARITY.
 FT DISULFID 48 61 BY SIMILARITY.
 FT DISULFID 51 69 BY SIMILARITY.
 FT DISULFID 72 87 BY SIMILARITY.
 FT DISULFID 93 113 BY SIMILARITY.
 FT DISULFID 115 128 BY SIMILARITY.
 FT DISULFID 134 152 BY SIMILARITY.
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 494 494 R -> K (IN REF. 2).
 SQ SEQUENCE 625 AA: 66621 MW: F8C1872E95511D8E CRK64;
 Query Match 13.9%; Score 312; DB 1; Length 625;
 Best Local Similarity 32.0%; Pred. No. 3.3e-11;
 Matches 79; Conservative 27; Mismatches 101; Indels 40; Gaps 7;
 Oy 18 VLGIFELLAASPP--QAVPYASSENTCRDQEKYEYRPHRRCSCSPRGTVASACSAI 75
 Db 15 LIALCVLPLDVTLLQVTPCTQE-----RHYEHNGR-CCSRCKGKYLSSCTPT 64
 Oy 76 RPTVCATCAENSYNEHMTLTICQLCPDPMVGLIEIAPRCSKRTQCRCPGMCAAM 135
 Db 65 SOSVCLPGCPDELDLTWNEDKCLLHKVDAGKALVAVDPGNHTARRCAGTAYH--W 121
 Oy 136 ALECHTCELLSDCPGETELKDEYKGNHCVCKAGHFQNTSPSARCOPTRCENOG 195
 Db 122 NSSCCECCRRNTFCAPFGAQHPQLQNLK-DTVCPTCLLGFSSDVSSTDRCKPTNCTLLG 180
 Oy 196 IVEAARGTQSDPTCKNPLPLEPPEMSSGLKRRRQSGEPNPVAGSMPEPKAHPPYPDV 255
 Db 196 IVEAARGTQSDPTCKNPLPLEPPEMSSGLKRRRQSGEPNPVAGSMPEPKAHPPYPDV 255

Db 181 KLEAHQGTESDVVC-----SSSMTLRRP-----PKEAQAAYLPSLI 216
 Oy 256 OPLUPTS 262
 Db 217 VLLIFTS 223
 RESULT 8
 TR1_HUMAN STANDARD; PRT; 616 AA.
 ID TR1_HUMAN
 AC Q9Y606;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Tumor necrosis factor receptor superfamily member 11A precursor
 DE (Receptor activator of NF-kB) (Osteoclast differentiation factor
 DE receptor) (ODFR).
 GN TNFRSF11A OR RANK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Dendritic cell;
 RX MEDLINE=98032977; PubMed=9367155;
 RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,
 RA Tomelesko M.E., Roux E.R., Teepe M.C., Dubose R.F., Cosman D.,
 RA Galibert L.;
 RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
 RT and dendritic-cell function.";
 RL Nature 390:175-179(1997).
 RN [2]
 RP MEDLINE=99097247; PubMed=9878548;
 RA Nakagawa N., Kinoshita M., Yamaguchi K., Shima N., Yasuda H., Yano K.,
 RA Morinaga T., Higashio K.;
 RT "RANK is the essential signaling receptor for osteoclast
 RT differentiation factor in osteoclastogenesis.";
 RL Biochem. Biophys. Res. Commun. 253:395-400(1998).
 RN [3]
 RP VARIANT PEO 16-L--L-21 DUPL, VARIANT PDB2 13-A--L-21 DUPL, AND VARIANT
 RP V-192.
 RX MEDLINE=20082806; PubMed=10615125;
 RA Hughes A.E., Ralston S.H., Marken J., Bell C., Macpherson H.,
 RA Wallace R.G.H., van Hul W., Whyte M.P., Nakatsuka K., Hoy L.,
 RA Anderson D.M.;
 RT "Mutations in TNFRSF11A, affecting the signal peptide of RANK, cause
 RT familial expansile osteolysis.";
 RL Nat. Genet. 24:45-48(2000).
 CC -1- FUNCTION: Receptor for TNFRSF11/RANKL/TRANCE/OPGL; essential for
 CC RANKL-mediated osteoclastogenesis. Involved in the regulation of
 CC interactions between T-cells and dendritic cells.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS EXPRESSION WITH HIGH LEVELS IN
 CC SKELETAL MUSCLE, THYMUS, LIVER, COLON, SMALL INTESTINE AND ADRENAL
 CC GLAND.
 CC -1- DISEASE: DEFECTS IN TNFRSF11A ARE THE CAUSE OF FAMILIAL EXPANSILE
 CC OSTEOLYSIS (PEO), A RARE AUTOSOMAL DOMINANT BONE DISORDER
 CC CHARACTERIZED BY FOCAL AREAS OF INCREASED BONE REMODELING. THE
 CC OSTEOCLYTIC LESIONS DEVELOP USUALLY IN THE LONG BONES DURING EARLY
 CC ADULTHOOD. PEO IS OFTEN ASSOCIATED WITH EARLY ONSET DEAFNESS AND
 CC LOSS OF DENTITION.
 CC -1- DISEASE: DEFECTS IN TNFRSF11A ARE A CAUSE OF FAMILIAL PAGET
 CC DISEASE OF BONE, ALSO KNOWN AS PAGET DISEASE OF BONE 2 (PDB2). IT
 CC IS A BONE REMODELING DISORDER WITH CLINICAL SIMILARITIES TO PEO.
 CC UNLIKE PEO, HOWEVER, AFFECTED INDIVIDUALS HAVE INVOLVEMENT OF THE
 CC AXIAL SKELETON WITH LESIONS IN THE SPINE, PELVIS AND SKULL.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

Query Match	Best Local Similarity	Score	DB 1:	Length	DB 2:	Indels	Gaps
Matches 92; Conservative	29.8%; Pred. No. 1.2e-11;	38; Mismatches 120;	Indels 59;	Gaps 14;			
16	PVLGLFG-LLAASOPAVPYASENOTCRDOEKYEPYEPQHRICCSRCPPGYTSAKCSR 74						
5	PLQCVLMGGLTNAVHE--PPTA-----CR--EKQYLINSQ--CSCLCOPGOKRIVSPCE 53						
75	IRDTVCATCAENSYNENHMYLTLCQLCRPPQVWGLEIACTSKRKTQRCORPGMFCMA 134						
54	FTFTECLPGGESEFIDTVMREHCHQHXYCDNLGLRYQCKGTSETPTTICBEGMHTS 113						
135	WALECTHCELLSCPPGTAEELKDEYKGNHCVCKAGHPONTSSSARCQPTTCENO 194						
114	EA--CESCVLHRCSSGPGVKK-QINATGVSDTICBPCPVGFSSNVSFAFEKCHPWTSCETK 170						
195	GLVEAAPAGTASDTTCKNPLEPLP-----PENSGL-----LKRPPGEGPNP 237						
171	DLVVOAQGTAKTIDYVC-GPQDRIRALVLPILFGLFALLVLVFIKKVAKKPPNKAPHP 229						
238	VAGSWEPKHPKPPPLVQPLPLISGDVSPVSTGLPAPVLEAGVPOQOSPLDLTREPOL 297						
230	---KQEPQELN--FPD-----DLGSTAPAVQETLHGCPVTO----- 263						
298	EPGEQSOVA 306						
264	EDGKESRIS 272						

CC from colon, stomach, rectum, esophagus and in SW480 colon
CC carcinoma cells.

CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: AF104419; AAD03056.1; -
DR EMBL: AF134240; AAD29688.1; -
DR EMBL: AF217796; AAF35244.1; -
DR EMBL: AF217793; AAF33685.1; -
DR EMBL: AF217794; AAF33686.1; -
DR EMBL: AL121845; CAC03668.1; -
DR EMBL: BC017065; AAH17065.1; -
DR Genew: HGNC:11921; TNFRSF6B.
DR MIM: 603361; -
DR HSSP: 014763; 1D0G.
DR Interpro: IPR001368; TNFR_C6.
DR Pfam: PF00020; TNFR_C6; 4.
DR ProDom: PD000771; TNFR_C6; 1.
DR SMART: SM00208; TNFR; 3.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KM Receptor: Apoptosis: Glycoprotein; Repeat: Signal.
FT SIGNAL 1 29
FT CHAIN 30 300
FT REPEAT 31 70
FT REPEAT 72 113
FT REPEAT 115 150
FT REPEAT 152 193
FT DISULFID 49 62
FT DISULFID 52 70
FT DISULFID 73 88
FT DISULFID 91 105
FT DISULFID 95 113
FT DISULFID 115 126
FT DISULFID 132 150
FT DISULFID 153 168
FT DISULFID 174 193
FT CARBOHYD 173 173
SQ SEQUENCE 300 AA: 32679 MW; F90AE33718449AF CRC64;
N-LINKED (GLCNAC...) (POTENTIAL).

Query Match 14.2%; Score 319; DB 1; Length 300;
Best Local Similarity 33.0%; Pred. No. 6; 9e-12;
Matches 88; Conservative 24; Mismatches 103; Indels 52; Gaps 12;

QY 6 ATSPAGLAMPVILGFLGLAASQPOAV--PPYASENQCROEKEYEPOHRCSCRC 62
DB 3 ALEGPILSLCLVLPALLPRAVGVAFETPY-----PWRDA-----ETGERLCAQC 52
QY 63 PPGTYVSACSRIRDTVCATCAENSYNEHWNLTITQLRCPDYPVNG--LEETAPCTSKR 120
DB 53 PPGTFVQRCRDRSPTTCGPCPRHYTQFWNYL---ERCRCYVNLGGEHEEARACHAVH 109
QY 121 KTGRCRQPMFC--AAMALCTHCELLSDCPGTEALKREVGKGNHCVPCRAQHQNQTS 179
DB 110 NRACRCRTGFFAHAGCLF--H---ASCPGA--GVIAQGTSPQNQCPCPGPSPASS 162
QY 180 SPFARCPTRCENGLVFAAPCTAOSDTCKN---PLEPDP----- 219
DB 163 SSEGQCPHRCATGALANVRSSSHDTLCSTGFPILSTRVGAEECEBRAVIDVARQ 222
QY 220 EMSGSLKR-----RPGEGENPVAG 240
DB 223 DISIKRLQRLQALFAPEGWGPFRAG 249

RESULT 6
ID TRN5_HUMAN
AC P25942; O9BYU0;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 5 precursor
DE (CD40L receptor) (B-cell surface antigen CD40) (CDW40) (Bp50).
GN TNFRSF5 OR CD40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM I).
RX MEDLINE=89356608; PubMed=2475341;
RT Stamenkovic I., Clark E.A., Seed B.;
RT "A B-lymphocyte activation molecule related to the nerve growth
RT factor receptor and induced by cytokines in carcinomas.";
RL EMBO J. 8:1403-1410(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM I).
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggaley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cordy N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Ellingham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckie E., Hunt A.R., Hunt S.E., Jekesch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levasialho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., Murray A.A.,
RA Milne S.A., Misty D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Philimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sultston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM II).
RX MEDLINE=21117110; PubMed=11172023;
RA Tone M., Tone Y., Fairchild P.J., Wykes M., Waldmann H.;
RT "Regulation of CD40 function by its isoforms generated through
RT alternative splicing.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:1751-1756(2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM I).
RC TISSUE=ovary;
RA Strusberg R.;
RL Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
RN [5]
RP 3D-STRUCTURE MODELING OF 24-144.
RX MEDLINE=97189482; PubMed=9037712;
RA Bajorath J., Aruffo A.;
RT "Construction and analysis of a detailed three-dimensional model of
RT the ligand binding domain of the human B cell receptor CD40.";
RL Proteins 27:59-70(1997).
RN [6]
RP 3D-STRUCTURE MODELING OF 26-186 IN COMPLEX WITH CD40L.
RX MEDLINE=98266353; PubMed=9605317;

RA Jacob C.O., Liu J.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-22 FROM N.A.
 RC TISSUE=Liver;
 RA Kissomerginis M., Fellowes R., Feldmann M., Chernajovsky Y.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Receptor with high affinity for TNFSF2/TNF-alpha and
 CC approximately 5-fold lower affinity for homotrimeric
 CC TNFSF1/Lymphotoxin-alpha (By similarity).
 CC SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.sib.ch).
 CC -----
 DR EMBL: M60469; AAA39752.1; -;
 DR EMBL: M59378; AAA40463.1; -;
 DR EMBL: U39488; AAA85021.1; -;
 DR EMBL: X87128; CAA60618.1; -;
 DR PIR: B38634; B38634;
 DR HSP: P19438; INCF.
 DR MGD: MGI:131483; Tnfslfb.
 DR InterPro: IPR01368; TNFR_c6.
 DR Pfam: PF00020; TNFR_c6; 4;
 DR PRODOM: PD000771; TNFR_c6; 1;
 DR SMART: SM00208; TNFR; 4;
 DR PROSITE: PS00652; TNFR_NGFR_1; 2;
 DR PROSITE: PS50050; TNFR_NGFR_2; 3;
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 474
 FT TUMOR NECROSIS FACTOR RECEPTOR
 FT SUPRAFAMILY MEMBER 1B.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT TNFR-CYS 1.
 FT TNFR-CYS 2.
 FT TNFR-CYS 3.
 FT REPEAT 120 164
 FT REPEAT 165 203
 FT TNFR-CYS 4.
 FT DISULFID 40 54
 FT BY SIMILARITY.
 FT DISULFID 55 68
 FT BY SIMILARITY.
 FT DISULFID 58 76
 FT BY SIMILARITY.
 FT DISULFID 79 94
 FT BY SIMILARITY.
 FT DISULFID 97 111
 FT BY SIMILARITY.
 FT DISULFID 101 119
 FT BY SIMILARITY.
 FT DISULFID 121 127
 FT BY SIMILARITY.
 FT DISULFID 136 145
 FT BY SIMILARITY.
 FT DISULFID 139 163
 FT BY SIMILARITY.
 FT DISULFID 166 181
 FT BY SIMILARITY.
 FT CARBOHYD 69
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 195
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 474 AA; 50319 MW; 462EAE398C4D6363 CRC64;
 Query Match 14.6%; Score 327.5; DB 1; Length 474;
 Best Local Similarity 29.5%; Pred. No. 3.4e-12;
 Matches 101; Conservative 40; Mismatches 142; Indels 59; Gaps 16;

DB 114 QONRVACGEAGRYCALKTHSGSCGRCMLSKCGPGF-GVASSRAPNGNVLCKACAPGTES 172
 QY 177 NTSSPSARCPHTRCENOGIVEAPGTAGSDTTCKNLEPLPEMGSLKRRPQGEQPN 236
 DB 173 DTSTSTVDCRPHRIS----ILAIQNASYDVA--PESPTLSAIPRTLVSOPEPTRSQ 227
 QY 237 PVAGSWEPRAHYFPPLVPL-----LPTSGVSPYSTGLPAPVLEAG 281
 DB 228 PL--DDEPQSGQT--PSILNLSSTPIIESTKGISLPIGLIVGYTSLGL-----LMLG 278
 QY 282 -----VPOQSPDLDTREPOLE--PGEOSVAHGNGTGHV 314
 DB 279 LVNCTILVQRKKRPSCLGRDAKVPHPYDEKSDQAVGLEQGH 320
 RESULT 5
 TR6B HUMAN STANDARD; PRT; 300 AA.
 ID TR6B_HUMAN
 AC 095407;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 6B precursor (Decoy
 DE receptor for Fas ligand) (Decoy receptor 3) (DCR3) (M68).
 GN TNFRSF6B OR DCR3 OR TR6.
 OS Homo sapiens (human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal lung;
 RX MEDLINE=99087326; PubMed=9872321;
 RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
 RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
 RA Goddard P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
 RA Goddard A.D., Botstein D., Ashkenazi A.;
 RT "Genomic amplification of a decoy receptor for Fas ligand in lung and
 RT colon cancer";
 RL Nature 396:699-703(1998).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-35.
 RC TISSUE=Prostate;
 RX MEDLINE=99233915; PubMed=10318773;
 RA Yu K.-Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
 RT "A newly identified member of tumor necrosis factor receptor
 RT superfamily (TR6) suppresses Listeria mediated apoptosis";
 RL J. Biol. Chem. 274:13733-13736(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=20122600; PubMed=10655513;
 RA Bal C., Connolly B., Metzger M.L., Hilliard C.A., Liu X., Sandly V.,
 RA Soderman A., Gallows S.M., Liu Q., Austin C.P., Caskey C.T.;
 RT "Overexpression of M68/DCR3 in human gastrointestinal tract tumors
 RT independent of gene amplification and its location in a four-gene
 RT cluster";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Mathews L.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Straubeberg R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Decoy receptor for the cytotoxic ligands TNFSF14/LIGHT
 CC and TNFSF6/Ra1. Protects against apoptosis.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC detected in adult stomach, spinal cord, lymph node, trachea,
 CC spleen, colon and lung. Highly expressed in several primary tumors


```

FT DOMAIN 245 415 CYTOPLASMIC (POTENTIAL).
FT REPEAT 42 81 TNFR-CYS 1.
FT REPEAT 82 124 TNFR-CYS 2.
FT REPEAT 125 170 TNFR-CYS 3.
FT REPEAT 171 213 TNFR-CYS 4.
FT DISULFID 43 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 80 BY SIMILARITY.
FT DISULFID 83 98 BY SIMILARITY.
FT DISULFID 101 116 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 126 132 BY SIMILARITY.
FT DISULFID 139 150 BY SIMILARITY.
FT DISULFID 142 169 BY SIMILARITY.
FT DISULFID 172 187 BY SIMILARITY.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 415 AA: 44956 MM: 298326566AEF61 CRC64;

```

```

Query Match 62.5%; Score 1404; DB 1; Length 415;
Best Local Similarity 63.9%; Pred. No. 1.8e-73;
Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

```

```

OY 1 MLPMATASAGLAMPVILGFLGILASOPQAVPVASENORODKEKEYEPOHRICCS 60
DB 1 MRLEPRASSPCGLAMGPIILGLSLVASOPQVLPYRIENOTCDQDKETEPFMDHYCCS 60
OY 61 RCPGTVYSAKSKSRIRDTVCATCAENSYNEHWNVLITICLRDPDPVAGLEIAPCTSKR 120
DB 61 RCPGEEVEFVAVCSRSQDTVCKTCOPHNSYNEHWNHLSYCOLCRPDIVLGEFEVAPCTSDR 120
OY 121 KTCRCRCPGMECAAMALECHTHC-ELLSDCPGTEALKEVNGGNNHCYPCAGHQRNT 178
DB 121 KAEBCRCRCPGSCVYLIDNECVCHCEERLYLQCPGTEAEVTEIDMTDVCVPCRGHQRNT 180
OY 179 SSPSARCOPTHRCENOGIVEAAPTASDPTKKNLEP----- 216
DB 181 SSPARCOPTHRCITGLVEAPRTSYSDITCKNPPRGAMLLAILLSLVLFLLFTTVL 240
OY 217 -----LPEPMS--GSLIKRRPGEGEPNVAGSNEPPKAMPYPDVLOPLIPISGDVSPV 268
DB 241 ACAMWRHPSLCRKLTGLTLKRHPGESESPCPA---PRADPHFDLEPLIPMSGDLSPS 296
OY 269 STGLPAPVLEAGVPOQOSPLDLTRPQLPEPESQVAGHNGIHVGGSTITGNITYIY 328
DB 297 PAGPPTASLEEVYLOQOSPLVQARELEAPGEHGVAGHNGIHVGGSTVIGNITYIY 356
OY 329 NGPVLAGPPEGDDLPATPEPPYPIPEEGDPPPGISLTPHODGKAMHLAETEHCG 383
DB 357 NGPVLAGTGRGDDPPAPPEPPYPIPEEGAPGPESELSTPYQDGDKAMHLAETETLG 411

```

```

RESULT 3
TRIB_HUMAN STANDARD; PRT; 461 AA.
AC P20333; O16042;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 1B precursor (Tumor
DE necrosis factor receptor 2) (P80) (TNF-R2) (p75) (CD120b) (Etarcept)
DE [Contains: Tumor necrosis factor binding protein 2 (TNFPII)].
GN TNFRSF1B OR TNFR2 OR TNFR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP MEDLINE=90260639; PubMed=2160731;
RA Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R.,
RA Dower S.K., Cosman D., Goodwin R.G.;
RT "A receptor for tumor necrosis factor defines an unusual family of

```

```

RT cellular and viral proteins.";
RL Science 248:1019-1023(1990).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91045991; PubMed=2172983;
RA Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,
RA Hale K.K., Squires C.H., Thompson R.C., Vanice J.L.;
RT "A second tumor necrosis factor receptor gene product can shed a
RT naturally occurring tumor necrosis factor inhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96299745; PubMed=8661109;
RA Beltinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,
RA Lepaslier D., Stallard B.J., Goeddel D.V., Desautels F.J.,
RA Broder G.M.;
RT "Physical mapping and genomic structure of the human TNFR2 gene.";
RL Genomics 35:94-100(1996).
[4]
RP SEQUENCE OF 37-461 FROM N.A.
RX MEDLINE=91370690; PubMed=1966549;
RA Demic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R.,
RA Brockhaus M., Lesslauer W.;
RT "Two human TNF receptors have similar extracellular, but distinct
RT intracellular, domain sequences.";
RL Cytokine 2:231-237(1990).
[5]
RP SEQUENCE OF 116-461 FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=90349572; PubMed=216946;
RA Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D.,
RA Ringold G.M.;
RT "Complementary DNA cloning of a receptor for tumor necrosis factor
RT and demonstration of a shed form of the receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).
[6]
RP SEQUENCE OF 27-31.
RX MEDLINE=90110215; PubMed=2153136;
RA Engelmann H., Novick D., Wallach D.;
RT "Two tumor necrosis factor-binding proteins purified from human
RT urine. Evidence for immunological cross-reactivity with cell surface
RT tumor necrosis factor receptors.";
RL J. Biol. Chem. 265:1531-1536(1990).
[7]
RP SEQUENCE OF 23-40; 65-69; 136-141; 300-306 AND 346-362.
RX MEDLINE=91056048; PubMed=2173696;
RA Loetscher H., Schlaeger E.J., Lahm H.W., Pan Y.-C.E., Lesslauer W.,
RA Brockhaus M.;
RT "Purification and partial amino acid sequence analysis of two
RT distinct tumor necrosis factor receptors from HU60 cells.";
RL J. Biol. Chem. 265:20131-20138(1990).
[8]
RP CHARACTERIZATION.
RX MEDLINE=93016040; PubMed=1328224;
RA Pennica D., Lam V.T., Mize N.K., Weber R.F., Lewis M., Fendly B.M.,
RA Lipari M.T., Goeddel D.V.;
RT "Biochemical properties of the 75-kDa tumor necrosis factor receptor.
RT Characterization of ligand binding, internalization, and receptor
RT phosphorylation.";
RL J. Biol. Chem. 267:21172-21178(1992).
[9]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN COMPLEX WITH
RP TNF2.
RX MEDLINE=99221490; PubMed=10206649;
RA Park Y.C., Burklit V., Villa A.R., Tong L., Wu H.;
RT "Structural basis for self-association and receptor recognition of
RT human TNF2.";
RL Nature 398:533-538(1999).
[10]
RP FUNCTION: Receptor with high affinity for TNFSF2/TNF-alpha and
RP approximately 10-fold lower affinity for homotrimeric
RP TNFSF1/lymphotoxin-alpha.
[11]
RP SUBCELLULAR LOCATION: Type I membrane protein and secreted.
RP -1- PTM: Phosphorylated; mainly on serine residues and with a very low
RP level on threonine residues.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)

DR EMBL: L04270; AAA6757.1; -
 DR EMBL: BC026262; AAH26262.1; -
 DR HSSP: P25942; ICDF.
 DR Genew: HGNC:6718; LTRB.
 DR MIM: 600979; -
 DR InterPro: IPR001368; TNFR_C6.
 DR Pfam: PF00020; TNFR_C6; 4.
 DR ProDom: PD000771; TNFR_C6; 1.
 DR SMART: SM00208; TNFR; 4.
 DR PROSITE: PS00652; TNFR_NGFR.1; 2.
 DR PROSITE: PS00505; TNFR_NGFR.2; 3.
 DR Receptor: Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 435
 FT TRANSMEM 31 227
 FT DOMAIN 228 248
 FT REPEAT 249 435
 FT REPEAT 42 81
 FT REPEAT 82 124
 FT REPEAT 125 168
 FT REPEAT 169 211
 FT DISULEFID 43 58
 FT DISULEFID 59 72
 FT DISULEFID 62 80
 FT DISULEFID 83 98
 FT DISULEFID 101 116
 FT DISULEFID 104 124
 FT DISULEFID 126 132
 FT DISULEFID 139 148
 FT DISULEFID 142 167
 FT DISULEFID 170 185
 FT CARBOHYD 40 40
 FT CARBOHYD 177 177
 SQ SEQUENCE 435 AA; 46709 MW; 6246266022F656F CRC64;
 Query Match 98.8%; Score 2220; DB 1; Length 435;
 Best Local Similarity 91.7%; Pred. No. 7.1e-120;
 Matches 399; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MLTPWATISAPGLAMGLVLTGLFGLAASOPQAVPYASENQTGRDQKEYYEPOHRIICGS 60
 DB 1 MLTPWATISAPGLAMGLVLTGLFGLAASOPQAVPYASENQTGRDQKEYYEPOHRIICGS 60
 QY RCPGTYVSAKSGRIDYVATCAENSYNHNWYLTICQLCRPCDPVYGLLETAIPCTSKR 120
 DB RCPGTYVSAKSGRIDYVATCAENSYNHNWYLTICQLCRPCDPVYGLLETAIPCTSKR 120
 QY 121 KTCRCQOPGMECAAMALECTHCELLSDCPGTAELKDEYKGNNNHCVPKAGHPONTSS 180
 DB 121 KTCRCQOPGMECAAMALECTHCELLSDCPGTAELKDEYKGNNNHCVPKAGHPONTSS 180
 QY 181 PSARQCPHTRCENOGIVEAPAGTASDTCKNLEPLPEPMS-----222
 DB 181 PSARQCPHTRCENOGIVEAPAGTASDTCKNLEPLPEPMS-----222
 QY 223 -----GSLKRRPQGGGPNPVAGSMPEPPRAHPYFPDLPVLPPLISGD 264
 DB 241 ANVEISCIWKSHPSLCKRLASLTKRRQGGGPNPVAGSMPEPPRAHPYFPDLPVLPPLISGD 300
 QY 265 VSPVSTGLPAAPVLEAGVPOQOSPLDLTREPQLEPEGSOVAGHTNGIHVTGSGMTITGN 324
 DB 301 VSPVSTGLPAAPVLEAGVPOQOSPLDLTREPQLEPEGSOVAGHTNGIHVTGSGMTITGN 360

QY 325 IYINPVLGPPGPDLPATPEPPYPIPEEGDPGISTPOHDEGKAMHIAETEHCA 384
 DB 361 IYINPVLGPPGPDLPATPEPPYPIPEEGDPGISTPOHDEGKAMHIAETEHCA 420
 QY 385 TFSNKGPRNQFTIHD 399
 DB 421 TFSNKGPRNQFTIHD 435

RESULT 2

TNR3_MOUSE STANDARD: PRI; 415 AA.
 AC P50284;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 3 precursor
 DE (Lymphotoxin-beta receptor).
 GN LTRB OR TNFRSF3 OR TNFR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CVB; TISSUE=Lung;
 RX MEDLINE=96072804; PubMed=7594541;
 RA Force W.R., Walter B.N., Hession C., Tizard R., Kozak C.A.,
 RA Browning J.L., Ware C.F.,
 RT "Mouse lymphotoxin-beta receptor. Molecular genetics, ligand binding,
 RT and expression."
 RL J. Immunol. 155:5280-5288(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96163885; PubMed=8586432;
 RA Nakamura T., Tashiro K., Nazarea M., Nakano T., Sasayama S.,
 RA Honjo T.;
 RT "The murine lymphotoxin-beta receptor cDNA: isolation by the signal
 RT sequence trap and chromosomal mapping."
 RL Genomics 30:312-319(1995).
 CC -1- FUNCTION: Receptor for the heterotrimeric lymphotoxin containing
 CC LTA and LTB, and for TNFS14/LIGHT. Promotes apoptosis via TRAF3
 CC and TRAF5. May play a role in the development of lymphoid organs
 CC (By similarity).
 CC -1- SUBUNIT: Self-associates (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)

DR EMBL: U29173; AAA68964.1; -
 DR EMBL: U38423; AAB00846.1; -
 DR EMBL: U30798; AAB1334.1; -
 DR HSSP: O14763; IDOG.
 DR MGD: MGI:104875; Ltbr.
 DR InterPro: IPR001368; TNFR_C6.
 DR Pfam: PF00020; TNFR_C6; 3.
 DR ProDom: PD000771; TNFR_C6; 1.
 DR SMART: SM00208; TNFR; 3.
 DR PROSITE: PS00652; TNFR_NGFR.1; 2.
 DR PROSITE: PS00505; TNFR_NGFR.2; 3.
 KW Receptor; Apoptosis; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 415
 FT DOMAIN 31 223
 FT TRANSMEM 224 244
 FT POTENTIAL.
 FT EXTRACELLULAR (POTENTIAL).

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:39:30 ; Search time 27.6158 Seconds

(without alignments)
\$99.261 Million cell updates/sec

Title: US-09-917-372-1

Perfect score: 2248

Sequence: 1 MLIPWATSAPGLAWGPIVLG.....EHGATPSNRGPRNQFTIHD 399

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2220	98.8	435	1	TNR3_HUMAN
2	1404	62.5	415	1	TNR3_MOUSE
3	349.5	15.5	461	1	TR1B_HUMAN
4	327.5	14.6	474	1	TR1B_MOUSE
5	319	14.2	300	1	TR6B_HUMAN
6	314.5	14.0	277	1	TNR5_HUMAN
7	312	13.9	625	1	TR11_MOUSE
8	298	13.3	616	1	TR11_HUMAN
9	253	11.3	289	1	TNR5_MOUSE
10	251.5	11.2	269	1	TNR5_BOVIN
11	244.5	10.9	461	1	TR1A_PIG
12	244	10.9	401	1	TR1B_MOUSE
13	243	10.8	655	1	TR21_MOUSE
14	240	10.7	283	1	TR14_HUMAN
15	239.5	10.7	401	1	TR1B_HUMAN
16	239	10.6	595	1	TNR8_HUMAN
17	239	10.6	655	1	TR21_HUMAN
18	235.5	10.5	401	1	TR1B_MOUSE
19	232	10.3	271	1	TNR4_MOUSE
20	231	10.3	471	1	TR1A_BOVIN
21	228	10.1	454	1	TR1A_MOUSE
22	226	10.1	272	1	TNR4_MOUSE
23	225	10.0	349	1	CRMB_CAMPS
24	223	9.9	498	1	TNR8_MOUSE
25	222.5	9.9	455	1	TR1A_HUMAN
26	212	9.4	323	1	TNR6_BOVIN
27	210	9.3	349	1	CRMB_VARV
28	209.5	9.3	277	1	TNR4_HUMAN
29	207	9.2	351	1	CRMB_COMPR
30	204	9.1	461	1	TR1A_MOUSE
31	203.5	9.1	255	1	TNR9_HUMAN
32	196.5	8.7	425	1	TR16_MOUSE
33	194	8.6	332	1	TNR6_PIG

34	191.5	8.5	493	1	TNR8_RAT	P97525 rattus norv
35	189	8.4	325	1	VT2_SEVKA	P25943 shope fibro
36	185.5	8.3	256	1	TNR9_MOUSE	P20334 mus musculu
37	185	8.2	440	1	TR1B_HUMAN	O14763 homo sapien
38	183.5	8.2	326	1	VT2_MTXVL	P25825 myxoma viru
39	183.5	8.2	336	1	TNR6_HUMAN	P25445 homo sapien
40	183.5	8.2	416	1	TR16_CHICK	P18519 gallus galli
41	175.5	7.8	417	1	TR16_MOUSE	O940W1 mus musculu
42	175.5	7.8	427	1	TR12_HUMAN	P08138 homo sapien
43	164	7.3	327	1	TNR6_MOUSE	O93038 h tumor nec
44	162.5	7.2	327	1	TNR6_MOUSE	P25446 mus musculu
45	149.5	6.7	180	1	TR22_MOUSE	O9er62 mus musculu

ALIGNMENTS

RESULT 1	ID	TNR3_HUMAN	STANDARD:	PRT:	435 AA.
AC	P36941:	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)				
DR	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Tumor necrosis factor receptor superfamily member 3 precursor				
DE	(Lymphotoxin-beta receptor) (Tumor necrosis factor receptor. 2 related protein) (Tumor necrosis factor C receptor).				
GN	LTBR OR TNFRSF3 OR TNFCR.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]	SEQUENCE FROM N.A.			
RP	TISSUE=Liver;				
RC	MEDLINE=93252381; PubMed=8486360;				
RX	BAENS M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;				
RA	"Construction and evaluation of a cDNA library of human 12p				
RT	transcribed sequences derived from a somatic cell hybrid."				
RL	Genomics 16:214-218(1993).				
RL	[2]	SEQUENCE FROM N.A.			
RP	TISSUE=Lung;				
RC	STRAUSBERG R.;				
RA	Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.				
RL	[3]	FUNCTION.			
RP	MEDLINE=94225209; PubMed=8171323;				
RX	Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,				
RA	Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;				
RT	"A lymphotoxin-beta-specific receptor."				
RL	Science 264:707-710(1994).				
RL	[4]	CHARACTERIZATION.			
RP	MEDLINE=99223511; PubMed=10207006;				
RX	Wu M.-Y., Wang P.-Y., Han S.-H., Hsieh S.-L.;				
RA	"The cytoplasmic domain of the lymphotoxin-beta receptor mediates cell				
RT	death in Hela cells."				
RL	J. Biol. Chem. 274:11868-11873(1999).				
RL	[5]	FUNCTION.			
RP	MEDLINE=20261554; PubMed=10799510;				
RX	Rooney I.A., Butovich K.D., Glase A.A., Borboroglu S., Benedict C.A.,				
RA	Whitbeck J.C., Cohen G.H., Eisenberg R.D., Ware C.F.;				
RT	"The lymphotoxin-beta receptor is necessary and sufficient for				
RL	light-mediated apoptosis of tumor cells."				
RL	J. Biol. Chem. 275:14307-14315(2000).				
CC	-1- FUNCTION: Receptor for the heterotrimeric lymphotoxin containing				
CC	LTA and LTB, and for TNFSF14/LIGHT. Promotes apoptosis via TRAF3				
CC	and TRAF5. May play a role in the development of lymphoid organs.				
CC	-1- SUBUNIT: Self-associates.				
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.				
CC	-1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.				

QY	1162	CGGGGGGGCTCATGACTATCACTGCGACACATTTACATCTACAAATGAGACCACTACTGGGGG	1221
Db	154396	CTGACGAGCTGTGAGTGTGCACCGGCAATATCTACATTTACAAATGGGCGCACTGCTGGGGG	154455
QY	1222	GACCAACGGGGTCTCTGGAGACTCTCCAGCTACCCCCGAACTCCATACCTATCCCATTTCCGAG	1281
Db	154456	GAACACGGGGCCCTGGAGACCTCCACGCTCCCCCTGAGCTCTCAATACCCGCTCCGAG	154515
QY	1282	AGGGGAGCCCTGACCCTCCCGGGCTCTGTACACCCCAACCAAGAAATGGCAAGGCTTGGC	1341
Db	154516	AGGGAGCGGCTGAGCCCTCTGTGAGTGTCTACACCTTACAGAGAAATGGCAAGCTTGGC	154575
QY	1342	ACCTAGCGGAGACAG--CACTGTGTGCCACACCCCTTAAACAGGGGCCCAAGAACAA	1400
Db	154576	ATCTGGCTGAGACAGACACTAAGGTGCCAAAGACTCTTGAC--GGGGCCAAAGGACCTA	154633
QY	1401	TTTATCACCACAGATGATGACTAGTGTGAAAGAGCAAGAAAGAGGGGGGCACAAAGGCAC	1460
Db	154634	TTAGCACCCAGCTGTGATGTGCATCTGGGAAAGCAAGAGAGAGGGGTGCA--GCAC	154691
QY	1461	CTTTTCCTTGAAGCTGCCCTGCCACAGTGGGATTTCAGAGGGGCTGATGAGGCCCGGG	1520
Db	154692	CTTTTTC--TGAGGCTGCTTACCAACACAGGACTCCAGAG--TTTGGCTGGGCTCTGA	154746
QY	1521	GAAGAGAGCCCTAAGGATTAAGGCTCAGACACTCTGAGAGCAGGTGGCACTGGCTG	1580
Db	154747	GATACAGAGCCCTAAGGAGACTTGGCTCAGACACTCT--TCAGAGAACACTAGCGG	154802
QY	1581	GGTACGGTGCCTCCACAGACTTCCCTACTGCTCTAGCAAACTGAGGCGTCCCGGA	1640
Db	154803	GGCACCGTGGCCGCG--CAGGCGCTTGACGACCGCTGAGAACTTGAAGCTTGTGGA	154861
QY	1641	GACCCACCCACCCCTGGGGCTGCTCAGCTCAGGCAACGGACAGGGCAGATGATTAACAAC	1700
Db	154862	GACCCACCACTCCCAAG-----GCTCAGACACAGGCAAGGCTTGTGACCCACAC	154913
QY	1701	TGTGCCCACTACAGCAAGCGCGACCGGACACAGGACCGAGAGGAGCCGCACAGGTC	1760
Db	154914	ATGGCTCTGTAGCGCTTAAACATAGGACAGAGAGGAGCTGGG-----CCAGGGTCA	154965
QY	1761	CTGTGAAGAGAGTCAACGGGCCCTTAAAGATTTCTGGTGCTCATCCCAAGCTTTACA	1820
Db	154966	CTTTCAGAGACATC--CGGGACCTTTGCAAGGCTTTTGTGTCTCATCCCAACTTTCAGA	155024
QY	1821	GACCTTTGGGGTTCCACACTCAGCTGAGCTGAGGTAGACCTGCATGAAGATTAATTT	1880
Db	155025	GGCCTTAGGGGCTTGTAC-----TGCGTGGCAAGATTAACCTGGGTGATACATATGTGC	155080

Search completed: April 15, 2003, 15:12:59
Job time : 4171.99 secs

Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Meheswari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,I.T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokkenwo,S., Oguh,M., Okunwo,G.,
Orangunye,N., Oriedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Fu,L.L., Qilles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojoudokan,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scheier,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Meinshock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 187998)

Morley,K.C.

Direct Submission

Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 187998)

Morley,K.C.

Direct Submission

Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GVSA

Center clone name: CH230-100N7

----- Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 133860 bases at least Q40

Consensus quality: 142697 bases at least Q30

Consensus quality: 148841 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 64 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1031:	contig of 1031 bp in length
1032	1131:	gap of unknown length
1132	2334:	contig of 1203 bp in length
2335	2434:	gap of unknown length
2435	3488:	contig of 1054 bp in length
3489	3588:	gap of unknown length
3589	4860:	contig of 1272 bp in length
4861	4960:	gap of unknown length
4961	6021:	contig of 1061 bp in length
6022	6121:	gap of unknown length
6122	7174:	contig of 1053 bp in length
7175	7274:	gap of unknown length
7275	8532:	contig of 1258 bp in length
8533	8632:	gap of unknown length
8633	9711:	contig of 1079 bp in length

9712	9811:	gap of unknown length
9812	11308:	contig of 1457 bp in length
11309	11408:	gap of unknown length
11409	12604:	contig of 1196 bp in length
12605	12704:	gap of unknown length
12705	14114:	contig of 1410 bp in length
14115	14214:	gap of unknown length
14215	16706:	contig of 2432 bp in length
16707	16806:	gap of unknown length
16807	17875:	contig of 1069 bp in length
17876	17975:	gap of unknown length
17976	19554:	contig of 1679 bp in length
19555	19754:	gap of unknown length
19755	21683:	contig of 1929 bp in length
21684	21783:	gap of unknown length
21784	22871:	contig of 1088 bp in length
22872	22972:	gap of unknown length
22972	24566:	contig of 1655 bp in length
24567	24766:	gap of unknown length
24767	26943:	contig of 2177 bp in length
26944	27043:	gap of unknown length
27044	29057:	contig of 2014 bp in length
29058	29157:	gap of unknown length
29158	30973:	contig of 1816 bp in length
30974	31073:	gap of unknown length
31074	32838:	contig of 1765 bp in length
32839	32838:	gap of unknown length
32939	34616:	contig of 1678 bp in length
34617	34716:	gap of unknown length
34717	37198:	contig of 2482 bp in length
37199	37298:	gap of unknown length
37299	39477:	contig of 2179 bp in length
39478	39577:	gap of unknown length
39578	41815:	contig of 2228 bp in length
41816	41915:	gap of unknown length
41916	44488:	contig of 2573 bp in length
44489	44588:	gap of unknown length
44589	46062:	contig of 1474 bp in length
46063	46162:	gap of unknown length
46163	48328:	contig of 2166 bp in length
48329	48428:	gap of unknown length
48429	51395:	contig of 2967 bp in length
51396	51495:	gap of unknown length
51496	54159:	contig of 2664 bp in length
54160	54259:	gap of unknown length
54260	56383:	contig of 2124 bp in length
56384	56483:	gap of unknown length
56484	58779:	contig of 2236 bp in length
58780	58879:	gap of unknown length
58880	60911:	contig of 2032 bp in length
60912	61011:	gap of unknown length
61012	63907:	contig of 2896 bp in length
63908	64007:	gap of unknown length
64008	66202:	contig of 2135 bp in length
66203	66302:	gap of unknown length
66303	69738:	contig of 3436 bp in length
69739	69838:	gap of unknown length
69839	72410:	contig of 2572 bp in length
72411	72510:	gap of unknown length
72511	76219:	contig of 3709 bp in length
76220	76319:	gap of unknown length
76320	79029:	contig of 2710 bp in length
79030	79129:	gap of unknown length
79130	83311:	

```

* 53463 55170: contig of 1708 bp in length
* 55171 55270: gap of unknown length
* 55271 55697: contig of 1427 bp in length
* 56698 56797: gap of unknown length
* 56798 58037: contig of 1240 bp in length
* 58038 58137: gap of unknown length
* 60362 60461: gap of unknown length
* 60462 62335: contig of 1874 bp in length
* 62336 62435: gap of unknown length
* 62436 64906: contig of 2471 bp in length
* 64907 65006: gap of unknown length
* 65007 67476: contig of 2469 bp in length
* 67476 67576: gap of unknown length
* 67576 69653: contig of 2078 bp in length
* 69654 69754: gap of unknown length
* 69754 73831: contig of 4077 bp in length
* 73831 73931: gap of unknown length
* 73931 77047: contig of 3117 bp in length
* 77048 77147: gap of unknown length
* 77148 80438: contig of 3291 bp in length
* 80439 80538: gap of unknown length
* 80539 83532: contig of 2994 bp in length
* 83533 83632: gap of unknown length
* 83633 85754: contig of 2122 bp in length
* 85755 85854: gap of unknown length
* 85855 89203: contig of 3349 bp in length
* 89204 92865: gap of unknown length
* 92866 92965: gap of unknown length
* 92966 95763: contig of 2798 bp in length
* 95764 95863: gap of unknown length
* 95864 99698: contig of 3835 bp in length
* 99699 104352: gap of unknown length
* 99799 104352: contig of 4554 bp in length
* 104353 108165: gap of unknown length
* 108166 108265: contig of 3713 bp in length
* 108266 111945: gap of unknown length
* 111946 112045: contig of 3680 bp in length
* 112046 116301: gap of unknown length
* 116302 116401: contig of 4256 bp in length
* 116402 120825: gap of unknown length
* 120826 120925: gap of unknown length

```

Query Match 14.28; Score 281.8; DB 2; Length 178228;
 Best Local Similarity 68.28; Pred. No. 4.6e-47;
 Matches 579; Conservative 0; Mismatches 232; Indels 38; Gaps 12;

```

QY 1135 CCCACGGTACCAATGGATTATGTCACCGGGGCTTATGACTATCATCTGGCAACATCT 1194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152113 CCACAGGTGCGAATGCGATTGACGCGAGGCTCTGTGACTGTACCGCAATATCT 152054

QY 1195 ACATCTCAATGACACAGTACTGTGGGGGACACCGGGTCTGTGAGACCTCCAGCTACCC 1254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152053 ACATATCAATGAGGCGAGTGTGGGGGAGACAGGGGCTGTGAGACCTCCAGCTACCC 151994

QY 1255 CCGAGACTCCATACCCCTATCCGAGAGGGGAGACCTGTGGCCCTCCCGGCTCTTACAC 1314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151993 CTGAGCTTCATACCCGACCTCCGAGAGGAGACGCGCTGCGCTCTGAGCTGTACAC 151934

QY 1315 CCCACCGAAGATGGCAAGGCTTGGACCTAGCGAGAGACAGAG-CACGTGTGGCCACA 1373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151933 CCAACCGAGAGATGGCAAGGCTTGGATCTGCTAGACAGAGACAGTATGAGGTGCAAG 151874

QY 1374 CCTCTTACAGGGGGCCCAAGAACCAATTATGACATGACATGAGTGTGAGAAAA 1433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151873 ACCTCTGAC--GGGGCCCAAGGACATATGACACCAATGCGTGAAGCATCTGGAAAA 151816

QY 1434 GCGAAGAAGAGGGGGGACACAAGGACCTTCTCCCTTGAAGGCTCCCTCCACAGTGGGA 1493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151815 GCAAGGAGGAGAGGGGTCCA--GCACCTCTTCC--TGAGGCTGCTCTACCAACAGG-- 151761

```

```

QY 1494 TTCACAGGGGCTGTAGTAGAGGCCCGGAGACAGAGCCCTTAAGGATTTAAGGCTACAGCA 1553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151760 -ACTACAGAGATTGGTGGCGCTGTGAGATACAGAGCCCTTAAGGACTTGGCTCAGACA 151702

QY 1554 CCTCTGAGACAGAGTGGGCGACTGTGGGTAGCGGTGCCCTCCACAGAGACTCTCCCTACTG 1613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151701 CCTCT---TCAGAGAGACATACAGCGGGGACCGCTGCCCGGC--CAGGCCCTGTGACCG 151647

QY 1614 CCTGACAAACCTGAGGCGCTCCGGGAGAGACCCACCCACCCCTTGGGCTGTACGCTCA 1673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151646 CCTGACAAAGCTTACAGACTTGTGTGACAGACCCACAGCTCCCAAG-----GCTCA 151595

QY 1674 GCGACGAGACAGGCAATGATATACCACTGTGCCCTACAGACAGCGGACCGAGACAC 1733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151594 GACACAGGCAAGGCTGTGTGACCCACCATGCTCTGTGAGCCCTTAACATATGACAGAGAG 151535

QY 1734 GGCACCGAGGAGGAGCGCCACACAGGTCACTGACAGAGACGTCAGGGCCCTTAAGGAT 1793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151534 GGAAGCTGG-----CCAGAGGCTCACTTCAAGGACATC--CGGGGACCTTCAAGGCT 151484

QY 1794 TCGTGTGCTCATCCCAAGCTTCAGAGACCCCTTGGGGGTTCACACTTACAGTGAAGT 1853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151483 TTGTGTGCTCATCCCAAGCTTCAGAGACCCCTTGGGGGTTCAGTGAAGT 151428

QY 1854 AGGTAGACCTGCATGATGAATTAATTAAGGAGAGAGCGCTCCCTCCCTCTAGAGA 1913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151427 AGATAGACCTGTGATGAATGATGTCGATGAGAGAGACCCCTCCCTCTAGAGA 151371

QY 1914 GGAAGGAGAGGAGGATTAATTAACATAGAGGGGCTGTGATGATTTCTTGAATGAGTGA 1973
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151370 GGGGAGAGAGAGACATGATGAAGAGGAGGAGGGGCTTCAATTAATTAATGATGAGAGA 151311

QY 1974 GAGTTTGG 1982
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151310 CAGTTTGG 151302

```

RESULT 14
 AC125909
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-100N7, *** SEQUENCING IN PROGRESS
 *** 64 unordered pieces.
 ACCESSION
 AC125909
 VERSION
 HTG: HTGS_PHASE1.
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus.
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS
 1 (bases 1 to 187998)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oshman,F.R., Allen,C.,
 Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbarta,J., Benton,J., Blmage,K., Blankenburg,K., Bonini,D.,
 Bouck,J., Bowie,S., Brieve,M., Brown,M., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dalhorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhardt,C., Edgar,D., Edwards,C.C., Elhai,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
 Homsl,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichteage,O., Lieu,C., Liu,C., Liu,J., Liu,W., Lounseged,H.,

REFERENCE
AUTHORS

Ratus. 1 (bases 1 to 178228)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbataia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buihay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,D.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B.,
 Homsl,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Krivovic,J., Kuresh,A., Landry,N., Lead,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
 Lozad,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maneshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogub,M., Okunodu,G.,
 Orquive,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
 Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wang,O.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,T., Zorrilla,S., Nelson,D.,
 Weinstein,G., Wu,Y.F., Zhou,T., Zorrilla,S., Nelson,D.,
 Weinstein,G., Wu,Y.F., Zhou,T., Zorrilla,S., Nelson,D.,
 Unpublished
 2 (bases 1 to 178228)
 Worley,K.C.
 Direct Submission
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

 Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: GZOW
 Center clone name: CH230-362C16

 Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 116771 bases at least Q40
 Consensus quality: 123909 bases at least Q30
 Consensus quality: 129462 bases at least Q20

***** NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 64 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1	1173:	contig of 1173 bp in length
*	1174	gap of unknown length
*	1273:	gap of unknown length
*	1274	contig of 1162 bp in length
*	2435:	gap of unknown length
*	2535:	gap of unknown length
*	2536	contig of 1032 bp in length
*	2536	gap of unknown length
*	3567	gap of unknown length
*	3568	contig of 1563 bp in length
*	3681	gap of unknown length
*	5230:	contig of 1402 bp in length
*	5330:	gap of unknown length
*	6732:	gap of unknown length
*	6832:	gap of unknown length
*	7990:	contig of 1158 bp in length
*	8090:	gap of unknown length
*	8091	contig of 1551 bp in length
*	9642	gap of unknown length
*	9741:	gap of unknown length
*	9742	contig of 1035 bp in length
*	10777	gap of unknown length
*	10877	contig of 1029 bp in length
*	11906:	gap of unknown length
*	12005:	gap of unknown length
*	12005:	gap of unknown length
*	13320:	contig of 1315 bp in length
*	13321	gap of unknown length
*	13421	contig of 1258 bp in length
*	14679	gap of unknown length
*	14779	contig of 1073 bp in length
*	15852	gap of unknown length
*	15852	contig of 1642 bp in length
*	17593:	gap of unknown length
*	17594	contig of 1218 bp in length
*	17694	gap of unknown length
*	18911:	contig of 1134 bp in length
*	19012	contig of 1134 bp in length
*	20145:	contig of 1134 bp in length
*	20146	gap of unknown length
*	20246	contig of 1311 bp in length
*	21357	gap of unknown length
*	21657	contig of 1059 bp in length
*	22715:	gap of unknown length
*	22716	contig of 1134 bp in length
*	23949:	contig of 1134 bp in length
*	24049:	gap of unknown length
*	25331:	contig of 1282 bp in length
*	25332	gap of unknown length
*	25431:	contig of 1186 bp in length
*	25432	gap of unknown length
*	26617:	contig of 1186 bp in length
*	26618	gap of unknown length
*	28435:	contig of 1718 bp in length
*	28436	gap of unknown length
*	28535:	gap of unknown length
*	30660:	contig of 2125 bp in length
*	30661	gap of unknown length
*	30760:	gap of unknown length
*	32263:	contig of 1503 bp in length
*	32264	gap of unknown length
*	34165:	contig of 1802 bp in length
*	34166	gap of unknown length
*	34265:	gap of unknown length
*	35302:	contig of 1037 bp in length
*	35302:	gap of unknown length
*	35402:	gap of unknown length
*	37282:	contig of 1880 bp in length
*	37283	gap of unknown length
*	39160:	contig of 1778 bp in length
*	39160:	gap of unknown length
*	39260:	gap of unknown length
*	39261	contig of 1620 bp in length
*	40880:	gap of unknown length
*	40881	contig of unknown length
*	42436:	contig of 1456 bp in length
*	42437	gap of unknown length
*	42537	contig of 1845 bp in length
*	44381:	gap of unknown length
*	44382	contig of 1910 bp in length
*	46391:	contig of 1910 bp in length
*	46392	gap of unknown length
*	46492	contig of 1918 bp in length
*	48410	gap of unknown length
*	48510	contig of 1422 bp in length
*	49932	gap of unknown length
*	50032	contig of 1340 bp in length
*	51372	gap of unknown length
*	51472	contig of 1891 bp in length
*	53362:	contig of 1891 bp in length
*	53363	gap of unknown length

Query Match 22.4%; Score 444.6; DB 6; Length 527;
 Best Local Similarity 88.1%; Pred. No. 1.1e-79;
 Matches 518; Conservative 0; Mismatches 9; Indels 61; Gaps 1;

```

QY 47 GGGAGCCCTGGAGGCCCCGGCTGGCCCTCCGGCCCTGGGGTGCACATCGGCCCTGAGT 106
    |||||
Db 1 GGGAGCCCTGGAGGCCCCGGCTGGCCCTCCGGCCCTGGGGTGCACATCGGCCCTGAGT 60
QY 107 CCCGTCACAGCTCTGGGGTGGGGGACCGCGCCACCGCTCCACAGACGTGGGGCTC 166
    |||||
Db 61 CCCGTCACAGCTCTGGGGTGGGGGACCGCGCCACCGCTCCACAGACGTGGGGCTC 89
QY 167 CTGCTCTCTCCAGGCCCCCAGCTTGTGCGCCGCTGGCGAGTGGCCGCTCTCTCT 226
    |||||
Db 90 -----GCCGCTGGCGAGTGGCCGCTCTCTCT 119
QY 227 GCGTGGGCGACCTCTGCGCCCGCGCTGGCTGGGGGCTCTGCTGGGGCTCTTGG 286
    |||||
Db 120 GCGTGGGCGACCTCTGCGCCCGCGCTGGCTGGGGGCTCTGCTGGGGCTCTTGG 179
QY 287 GCTCTGGCAGCATCGACGCCCGCAGCGGCTGCTCATATGCTGGGAGAACAGACCTG 346
    |||||
Db 180 GCTCTGGCAGCATCGACGCCCGCAGCGGCTGCTCATATGCTGGGAGAACAGACCTG 239
QY 347 CAGGACCCAGAGAAAAGAACTATGAGCCCGACACCGCATCTGCTGCTCCGCTGCC 406
    |||||
Db 240 CAGGACCCAGAGAAAAGAACTATGAGCCCGACACCGCATCTGCTGCTCCGCTGCC 299
QY 407 GCGAGGCACTATGCTGAGTAAATGAGCGGATCGGGACACAGTTTGTGCCACATG 466
    |||||
Db 300 GCGAGGCACTATGCTGAGTAAATGAGCGGATCGGGACACAGTTTGTGCCACATG 359
QY 467 TGCCGAGATTCCTACACAGCAGCACTGAACTACCTGACCTGCGAGCTGCGGCC 526
    |||||
Db 360 TGCCGAGATTCCTACACAGCAGCACTGAACTACCTGACCTGCGAGCTGCGGCC 419
QY 527 CTGAGACCCAGTATGAGGCGCTCGAGAGATTTGCCCTGCAACAAGACGAGACCA 586
    |||||
Db 420 CTGAGACCCAGTATGAGGCGCTCGAGAGATTTGCCCTGCAACAAGACGAGACCA 479
QY 587 GTGCGCGTGGCAGCGCGGAAATGTTCTGTGCGCTGGGGCTCGAGAGT 634
    |||||
Db 480 GTGCGCGTGGCAGCGCGGAAATGTTCTGTGCGCTGGGGCTCGAGAGT 527

```

RESULT 12
 G11431/c 400 bp DNA linear SRS 30-MAR-2000
 LOCUS SHGC-10552 Human Homo sapiens SRS genomic, sequence tagged site.
 DEFINITION G11431
 ACCESSION G11431.1 GI:1017523
 VERSION SRS.
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 400)
 OLIVIER, M. and COX, D. R.
 Unpublished (2000)
 JOURNAL Unpublished (2000)

Contact: Michael Olivier, David R. Cox
 Stanford Human Genome Center
 Stanford University School of Medicine
 4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
 Tel: (650) 320-5800
 Fax: (650) 320-5801
 Email: oliviereshgc.stanford.edu
 Primer A: AGGAATCTTACCAACCC
 Primer B: GCGTCCACACATTCACGTG
 SRS size: 132
 PCR Profile:
 Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
 Annealing: 62 degrees C for 23 seconds
 Polymerization: 72 degrees C for 30 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600

Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/ul
 Total Vol: 10 ul

Buffer:
 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 20 mM
 pH: 8.3

FEATURES
 source Prepared with primer pairs derived from T62568 -- Merck/UniEST.
 Location/Qualifiers
 1..400
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="12"
 /clone_lib="Human"

SRS
 primer_bind 89..220
 primer_bind 89..107
 BASE COUNT complement(202..220)
 ORIGIN 74 a 118 c 104 g 103 t 1 others

Query Match 15.0%; Score 296.8; DB 11; Length 400;
 Best Local Similarity 97.0%; Pred. No. 8.e-50;
 Matches 323; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

```

QY 1652 CCCCTGGGGGCTGCTACGCTCAGGACGACGAGGGGACATGTA-CGAACTGCTG-CCCA 1709
    |||||
Db 400 CCCCTGGGGGCTGCTACGCTCAGGACGACGAGGGGACATGTA-CGAACTGCTG-CCCA 341
QY 1710 CTACAGCAGCGCGCAGCAGGAGCAGCGACCGAGGAGCGGCACAGGGTCACTGCAAG 1769
    |||||
Db 340 CTAAGCAGCGCGCAGCAGGAGCAGCGACCGAGGAGCGGCACAGGGTCACTGCAAG 281
QY 1770 ACCTACGCGCGCGCTCTAAGATTCGTGCTCATATCCCAAGCTTCAGAGACCTTTG 1829
    |||||
Db 280 ACCTACGCGCGCGCTCTAAGATTCGTGCTCATATCCCAAGCTTCAGAGACCTTTG 221
QY 1830 GGGTTCACACTTCACTGAGTGAAGTGAACCCGATGAAGATGAATTAATAGAGAGG 1889
    |||||
Db 220 GGGTTCACACTTCACTGAGTGAAGTGAACCCGATGAAGATGAATTAATAGAGAGG 161
QY 1890 ACGTCTCTTCCCTCCCTCTCTAGAGAGAGAGAGAGAGATTAACAACATAGAGGGGTTG 1949
    |||||
Db 160 ACGTCTCTTCCCTCCCTCTCTAGAGAGAGAGAGAGAGATTAACAACATAGAGGGGTTG 101
QY 1950 GGTAGATTCTCAGGTATGGGGAAGAGTTTGG 1982
    |||||
Db 100 GGTAGATTCTCAGGTATGGGGAAGAGTTTGG 68

```

RESULT 13
 AC128082 178228 bp DNA linear HTG 19-JUL-2002
 LOCUS AC128082/c
 DEFINITION Rattus norvegicus clone CH230-362C16, *** SEQUENCING IN PROGRESS
 ACCESSION AC128082
 VERSION AC128082.1 GI:21908679
 KEYWORDS HTG; HTGS PHASE1.
 SOURCE Rattus norvegicus.
 ORGANISM Rattus norvegicus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

```

PCDVLGFEVPACTSDRKABRCOPGMSVYLNDNEVHCHEERLVLCOPTGEAEVTD
EIMDYNVCVKRKHGFIQNTSSPRAROPHRCJGIVLAEARHGSYSDTICKNPPR
GAMLLILILSLVLEFETVLACAMHPSGLCRIGLILRHRGEPSPCPAPRAD
PHRPDLAPRLLPMGSDLSRSPAPRPTAPSLSEVLYLOOQSPVVOARELEAEFEGHOVA
HGANGIHVIGSVTVIGNITLYINGPVLGRTGPDPPAPPEPPYPTPEBGAPESELS
TPYEDGAMHLALETETLGCODL*
BASE COUNT      316 a      537 c      455 g      306 t
ORIGIN

Query Match      34.0% Score 673.8 DB 10; Length 1614;
Best Local Similarity 69.5%; Pred. No. 4.7e-126;
Matches 1111; Conservative 0; Mismatches 377; Indels 111; Gaps 10;

OY 48 GGAGCCCTGGAGCCCGGCGCTGCGCTCCGCGCCCTGGGAGTGCACATCGCCCTGAGTC 107
DB 25 GGAGCCCTGGAGCCCGGCGCGCTCCGCGCCCGGCGGGAAGCTGG----- 75
OY 108 CCGTCCAGAGCTCTGGGCTCGGGCAGCGCCGACCGCTGCGCCAGAGCT--CGGGCTTC 166
DB 76 ----CCAGCCCGCAGGTTTGGGAAGTCGTGGCAACGCTGCTCAGGAGCTCGGGCTTC 131
OY 167 CTGCTTCCTCCAGGCGCCCA--CGTTGCTGGCGCGCTGGCGAGTGGCGGCATGCTC 224
DB 132 CCACCTTCTCTAGACTCCACCGGCTGTGTCAGCCGAGCCGAAAGCGCCCATGCGC 191
OY 225 CTGCTTGGGCGACCTCTGCCCCGCGCTGGCTGGGCGCTCTGGTGGGCTCTTC 284
DB 192 CTGCCCCGGGCTCTGCTCTGCGGCGCTGGGCGGCGACTGCTGTGGAGCTCAGC 251
OY 285 GGGCTCTGGCAGCATCGAGCCCGCAGCGGCTGCTCATATGCTGGAGAACAGCAGCC 344
DB 252 GGGCTTGGGCGCTCTGAGCCCGCAGCGGCTGCTCATATGAGTAAAGAACAGCT 311
OY 345 TCGAGGAGCAGGAAAGATATGATGAGCCCGCAGCAGCGGATGCTGCTCCGCGTGC 404
DB 312 TGGTGGGAGCAGGAGAAATATGAGCCCGCAGCAGCGGATGCTGCTCCGCGTGC 371
OY 405 CCGCAGGACCTATGCTCAGCTAAATGATAGCCGATCGGAGACAGATTTGTGCCACA 464
DB 372 CCGCAGGAGGAGTGTGCTTGTGCGATGACCGCGCAGCAAGACAGGTTTGAAGACT 431
OY 465 TGTGCGAGAAATTCATACAGGAGCAGTGAACCTGACATCGGAGTGGAGTGGCGCG 524
DB 432 TCGCCCGCAATATTCATATATGAGAGTGAACCATCTCCACCGCGCAGCTGTGCCG 491
OY 525 CCTGTGACCCAGTATGGGCTCGAGAGATTTGCCCTCGACAGCAAGCAAGGAGGAGC 584
DB 492 CCTGTGACATTTGTGTGGGCTTGAAGAGTTGGCCCTTGCACAGCATGGGAAAGCC 551
OY 585 CAGTCCCGCTGCCAGCGGGAATGTTCTGTGCTGCTGGGCGCTCGATGTAACACTG- 643
DB 552 GAGTGGCGGTGTCAGCGGAGGATGTCGTGTATCTGGAATGAGTGTGCACTGT 611
OY 644 -----CGAGCTACTTTTACAGCCCGCTGGCAGTGAAGCCGAGTCAAAATGAGTT 698
DB 612 GAGGAGAGCGGCTGTACTGTGCGAGCTGGCAGAGGAGGCTGACATGAAAT 671
OY 699 GGAAGAGTAAACCACTGCTCCCTGCAAGGAGCGGACATTCAGAAATCTCTCC 758
DB 672 ATGAGTACTAGAGCTACAGTGTCTCCCTTAAGCCGGGACATTCAGAACTTCTCC 731
OY 759 CCGAGCGCGCGTGGCAGCGCCAGCAGAGTGTGAGAACCAAGTCTGGTGAAGGAGCT 818
DB 732 CCGTACGCGCGCTGTCACCCATACAGATGAGATCCAGGCGCTGTGTGAGGAGCT 791
OY 819 CAGAGCAGTGGCCAGTCCGACACAACTGCAAAATCCATTAGAGCACTGCCCGCAGG 878
DB 792 CAGAGTACTCTTACTCGGATACCATCTGTAATAATCCCGCAGAGCCAGGACCAATGCTA 851
OY 879 ATGTCAGATTCGCTCTCAG-----AGGCGTCC 907
DB 852 CTGCTGGCATCTGCTGTGCGTGTCTTCTGCTCTTCAACACACTGTCTGCGCTGT 911

```

```

OY 908 GCAGGAGAGGAGACCAATCTGTAGTGGAGCTGGAGC----- 948
DB 912 GCCTGGATAGAGCACCCTGCTCTGTGAGAAAGCTGGAGTCTGTCTCAAGGGGACCA 971
OY 949 -----CTCCGAAGGCCCATTCATCTTCCCTGACTTG 980
DB 972 GAGGAGAAAGATCTCCCTTGGCCCGCTCCAAAGCGGACCAACATTCCTGACCTG 1031
OY 981 GTACAGCCACCTGCTACCAATTTCTGGAGATGTTTCCCGATATTCAGTGGGCTCCCGCA 1040
DB 1032 GAGAGCCACCTCTCCATCTGTGTGAGACTTGTCCCTGCTGCTGGAGCCCGCAAG 1091
OY 1041 GCCCAGATTTTGGAGCAGAGGGGTGCGGCAACAGCAGAGTCTCTGAGCTGACAGGAG 1100
DB 1092 GCCCCTCTTGGAGAAATGAGTGTACAAACAGAGAGTCCCTGTGTCAGGCGAGAG 1151
OY 1101 CCGCAGTTGGAACCCGGGAGAGAGCCAGGTGGGCCAACAGTACCAATGGCAATTCATGTC 1160
DB 1152 CTGAGAGCTGAACCTGGGGAMCATGCGCAGGTGGCCCAAGGTGCGAATGGCATTCATGTG 1211
OY 1161 ACCGGGGGTCTATGACTATCACTGCAACATCTACATATGACATGAGCACTACTGGG 1220
DB 1212 ACCGGGGGTCTGTGTACTGTACCGGCAACATCTATATATGAGGCACTGCTGGGG 1271
OY 1221 GGACCAACCGGCTCTGAGAGACCTCCAGCTACCCCGAAGCTCCATACCCCATCCGAA 1280
DB 1272 GGAACACAGGGGGCTGAGAGACCTCCAGCTCCCTGAGCCTTCATACCCCATCTGAA 1331
OY 1281 GAGGGGAGCCTTGGGCTTCCGGGCTCTTACACCCCAAGAGAGATGGCAAGCTTGG 1340
DB 1332 GAGGGAGCTCTGCGGCTCTGAGCTGTCTACACCTTACAGAGAGATGGCAAGCTTGG 1391
OY 1341 CACCTAGCGGAGACAGAG--CACTGTGTGTCACACCTCTAACAGGGGCGCCAAAGAAC 1399
DB 1392 CATCTGGCGGAGACAGAGACATAGATGTCGAAGCTTGTGACAGG--CAAGGAACT 1449
OY 1400 ATTATACCCCATGACTGACTAGTGTGAGAAAAGGAGAGAGAGGGGCGCAAGGGCA 1459
DB 1450 ATTTGGCACCCCTGCTGAGGCGACCTGGGAAAGGACAGAGGGTCCAGCACCTT 1509
OY 1460 CCTTCTCCTTGAAGGCTGCGTGCACAGTGGGATTCACAGGGGCGCTGATAGGGCGCG 1519
DB 1510 CTTT-----TGAGCTGCTCTACTACCAAG--ACTACAGAGGCTGTGGAGGCTCTG 1561
OY 1520 GGAAGCAGAGCCCTAAGGATTAAAGCTCAGACACTCT 1558
DB 1562 AGAAATAGAGCCCTAAGGACTTAAAGCTCAGACACTCT 1600

RESULT 11
AX381072      527 bp      DNA      Linear      PAT 18-MAR-2002
LOCUS      AX381072
DEFINITION      Sequence 10 from Patent W00212280.
ACCESSION      AX381072
VERSION      AX381072.1 GI:19575891
KEYWORDS
SOURCE
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Pyle,R.A., Xu,J. and Secrist,H.
AUTHORS      Compositions and methods for the therapy and diagnosis of colon
TITLE      cancer
JOURNAL      Patent: WO 0212280-A 10 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
1..527
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      86 a      190 c      158 g      88 t      5 others
ORIGIN

```

QY 397 CCGCTCCCGCCAGGACCATATGCTCAGCTTAATGTAGCCGCATCCGGGACAGTTT 456
 Db 348 CCGCTGTCCCGCCAGGAGGAGTTTGTCTTTCGGGTATCCACCACCAGCAAGACGTTT 407
 QY 457 GTGCAATATGTGCGAATTTCTTCTACACAGCAGCTGGAATCTGACCATCTGCGAG 516
 Db 408 GCAAGACTTCCCGCCATTAATTTCTATATGATACCTGGAACCATCTCTCCACTGCGAG 467
 QY 517 TGTGCGCCCTGTGACCCAGTATGGGCTCGAGAGAGTGGCCCTGACAGCAAGCAAC 576
 Db 468 TGTGCGCCCTGTGACATTTGTCTGTGCTTATAGAGGTTGGCCCTGTGACAGGATC 527
 QY 577 GGAAGACCCAGTCCCGCTGCGACCGGGGATTTCTGTGCTGCTGCGCCCTGAGTGA 636
 Db 528 GGAAGACCCAGTCCCGCTGCGACCGGGGATTTCTGTGATCTGACATGAGTGTG 587
 QY 637 CACACTG-----CGAGTACTTTGACTGCGCCGCTGGACTGTAAGCCGAGCTCAAG 690
 Db 588 TGTACTGTGAGAGAGGAGGAGGCTTGTACTGTGCGACGCTGCGACAGAGGCTCAG 647
 QY 691 ATGAAGTTGGAGAGGAGTAAACCAACATGCGTCCCTGCAAGGAGGAGGCTTCCAGAA 750
 Db 648 ATGAATTTATGATGACTGACGTCATCTGTGCTGCTGTAAGCCGGAGACCTTCCAGAA 707
 QY 751 CCGTCCCGCCCGAGCGCCGCTGCGACGCCACACAGAGTGTGAGAACCAAGGCTGTG 810
 Db 708 CTTCCTCCCGCTGAGCGCCGCTGCAACCCATACAGAGTGTGAGATTCAGGCGCTG 767
 QY 811 AGGAGCTCCAGGACCTGCGCTGCGACAGCAACCTGCAAAATTCATTAGAGCCACTG 870
 Db 768 AGGAGCTCCAGGACCTGCTCTCTGATGATCATCTGTAATAATCCCGAGGCGAGAG 827
 QY 871 CCCCAGAGATGTCAGAGATGCTGCTCAAG----- 899
 Db 828 CATGCTACTGCTGCGCACTCTGCTGTGCTGCTCTTCTGTCTTCAACACTGTC 887
 QY 900 AGGCTCCCGAGGAGAGGAGGAGCCCAATCTGATGAGAGCTGGAGC----- 948
 Db 888 TGGCTGTGCTGAGTGAAGGACCCGCTCTCTGCAAGAACTGGTACTGCTGCTCAAGC 947
 QY 949 -----CTCGAAGGCCCATCATCTATCTCC 972
 Db 948 GGCACCCAGAGGAGAAATCTCCCTTCCCGCCGCTCCAGAGCCGACCACTTCC 1007
 QY 973 CTGACTTGTACAGCCACTGCTACCATTTCTGTGAGATGTTTCCCACTATCCATGGGC 1032
 Db 1008 CTGACTGTGAGAGCCACTCTACCATGCTGTGAGACTTGTCCCAATCCCTGCTGAC 1067
 QY 1033 TCCCGCAGCCCGAGTTTGTGAGGAGGAGGCGCGCAAGCAGAGTCTGTGACCTGA 1092
 Db 1068 CCCCAGCGCCCTCTCTGTGAGAGAGGAGTGTACAAAGCAAGTCCCTGCTCCAGG 1127
 QY 1093 CCAAGAGCCGAGTTGGAACCCCGAGGAGAGAGAGCTGGCCAGAGTACCAAGGCA 1152
 Db 1128 CCAAGAGCTGAGGAGTGAACCTGGGAGAAATGCGCAGGAGGCGCAAGGTCGAATGCA 1187
 QY 1153 TTTACTGACCGGCGGCTATGATCTATCTACCTGCAACATCTACATCTACATGAGCAG 1212
 Db 1188 TTTACTGACCGGCGGCTATGATCTATCTACCTGCAACATCTACATCTACATGAGCAG 1247
 QY 1213 TACTGAGGAGGAGCAGCGGCTGTGAGAGCTCCAGAGTATCCCGGAACTCCATACCCCA 1272
 Db 1248 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1307
 QY 1273 TTTCCCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1332
 Db 1308 CTCTCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1367
 QY 1333 AGGCTTGGCAGCTAGCGGAGAGAGAG--CACTGTGTGTGCAACACCTCTTACAGGGGCCCA 1391
 Db 1368 AAGCTTGGCAGCTAGCGGAG 1425

QY 1392 AGGAACCAATTTATACCATATGACTGACTGTGAGAAAAGGAGAGAGGGGCA 1451
 Db 1426 AGGAACCAATTTATACCATATGACTGACTGTGAGAGGAGAGAGAGAGAGAGAGAGAG 1485
 QY 1452 CAAGGAGCAGCTTCTCCCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1511
 Db 1486 AAGCAGCTTCTTTC-----TGAGGCTGTCTACTGACCAAG--ACTACAGGAGGCTGGGG 1537
 QY 1512 GGGCCCGGGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1568
 Db 1538 AGGCTGTGAGAAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1597
 QY 1569 GGGCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1628
 Db 1598 GAGCAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1656
 QY 1629 GGGCTCCCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1688
 Db 1657 GAGCAGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1716
 QY 1689 CATGATACCAACTGCTGCGCACTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1747
 Db 1717 AAGGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1774
 QY 1748 GCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1807
 Db 1775 GGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1833
 QY 1808 CCCAAGCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1867
 Db 1834 CCGTACTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1889
 QY 1868 TGAAGATGAATTTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1927
 Db 1890 GGAAGGTGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1949
 QY 1928 GTCAATTAACAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1962
 Db 1950 GAGAGGAGAG--TGGGCGCTTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2003

RESULT 10

MUSLYMPHOB

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

FEATURES

source

CDS

Location/Qualifiers

1..1614

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_line="Str-2"

186..1433

/note="putative"

/product="Lymphotoxin-beta receptor"

/protein_id="AA80846.1"

/db_xref="GI:600223"

/translation="MRPRASSPFGIAGMGLLGLSLIVASOPOLVPPYRIENQTCM
DQDEYEYPMHIVCCSRCPGPEVFVAVCSRSODTVCKTPHNSYNEHMHNLSTQICR

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1 Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
AUTHORS
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2135 11-APR-2002;
GENE LOGIC INC (US)

FEATURES
Location/Qualifiers
1. 2136
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 446 a 706 c 608 g 376 t
ORIGIN

Query Match Best Local Similarity 94.6%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

```

QY 51 GCCCTGAGAGCCCGGCTGCGCTCCGCGCTCCGCGCTGGGCGTGCATCGGGCCCTGAGTCCCG 110
Db 1 GCCCTGAGAGCCCGGCTGCGCTCCGCGCTCCGCGCTGGGCGTGCATCGGGCCCTGAGTCCCG 60
QY 111 TCCGAGGCTCTGGGCTCGGGCAGCCGCCGACCGCTGCCGAGAGCTCGGGCTCTCTGC 170
Db 61 TCCGAGGCTCTGGGCTCGGGCAGCCGCCGACCGCTGCCGAGAGCTCGGGCTCTCTGC 120
QY 171 CTTCCTCCAGGCCCGCCAGCTGCTGGCGGCTGGCGAGTGGCGCGCCATGCTCTGCT 230
Db 121 CTTCCTCCAGGCCCGCCAGCTGCTGGCGGCTGGCGAGTGGCGCGCCATGCTCTGCT 180
QY 231 TGGGCGACCTCTGCCCCCGGCTGGCGTGGGGGCTCTGTGTCTGGGCTCTTTCGGGCTC 290
Db 181 TGGGCGACCTCTGCCCCCGGCTGGCGTGGGGGCTCTGTGTCTGGGCTCTTTCGGGCTC 240
QY 291 CTGGCAGATGCGAGGCCCGCCAGCGGTGCTTCATATGCTGTGGAGAACAGACCTGCAGG 350
Db 241 CTGGCAGATGCGAGGCCCGCCAGCGGTGCTTCATATGCTGTGGAGAACAGACCTGCAGG 300
QY 351 GACCAAGAAAGAAATACATATAGAGCCCGACCGCATCTCTCTCCCGTGGCCGCA 410
Db 301 GACCAAGAAAGAAATACATATAGAGCCCGACCGCATCTCTCTCCCGTGGCCGCA 360
QY 411 GGCACATATGCTCAGCAATATGAGCGCATCCGGGACACAGTTTGTGCACATGTGCC 470
Db 361 GGCACATATGCTCAGCAATATGAGCGCATCCGGGACACAGTTTGTGCACATGTGCC 420
QY 471 GAGAAATCTCAACAGACATGGAATACCTGACATGCGCGAGCTGTGCCGCTGT 530
Db 421 GAGAAATCTCAACAGACATGGAATACCTGACATGCGCGAGCTGTGCCGCTGT 480
QY 531 GACCCAGTATGGGCTCGAGAGATTTGCCCTGCAAGCAAGCAAGCAAGCAAGCAAGTGC 590
Db 481 GACCCAGTATGGGCTCGAGAGATTTGCCCTGCAAGCAAGCAAGCAAGCAAGCAAGTGC 540
QY 591 CGCTGCAGCGGGAATGTTCTGTGCTGCTGGGCGCTCGAGTGTACACTGCGAGCTA 650
Db 541 CGCTGCAGCGGGAATGTTCTGTGCTGCTGGGCGCTCGAGTGTACACTGCGAGCTA 600
QY 651 CTTTCTGACTGCGGCTGCTGCTGAGAGCCGAGCTCAAAAGATGAAGTTGGAGAGGTAA 710
Db 601 CTTTCTGACTGCGGCTGCTGCTGAGAGCCGAGCTCAAAAGATGAAGTTGGAGAGGTAA 660
QY 711 AACCACTGCTGCGCTGCAAGGAGGCACTTCCAGAAATACCTCTCCCGCAGCGCCG 770
Db 661 AACCACTGCTGCGCTGCAAGGAGGCACTTCCAGAAATACCTCTCCCGCAGCGCCG 720
QY 771 TGCAGCGCCACACAGGTGTGAGAACCAAGGTCTGTGGAGGCACTCCAGGCACTGCG 830
Db 721 TGCAGCGCCACACAGGTGTGAGAACCAAGGTCTGTGGAGGCACTCCAGGCACTGCG 780
QY 831 CAATCCGACCAACTGCAAAAATTCATTAGAGCACTGCCCGCAGAGATGTCA----- 884

```

```

Db 781 CAGTCCGACACAACTGCAGAAAATTCATTAGAGCACTGCCCGCAGAGATGTCAAGAAC 840
QY 885 ----- 884
Db 841 ATGCGATGCTGGCGGTTCTGTGCTGCACTGGCCTTCTTTCGCTCTGCGACCGCTTTC 900
QY 885 -----GATCGCTGCTCAAGAG 902
Db 901 TCTTCGATCTGGAAGAGCCACCTCTTCTCTGCAAGAAACTGGATGCTCTCAAGAG 960
QY 903 CGTCCGAGGAGAGGAGCCCAATCTCTTACTGGAAGCTGGAGCTTCGAGAGGCCAT 962
Db 961 GCTCCGAGGAGAGGAGCCCAATCTCTTACTGGAAGCTGGAGCTTCGAGAGGCCAT 1020
QY 963 CCATATCTCCCTGACTTGTGTACAGCCACTGCTACCCATTTTGTGAGATGTTTCCCGATA 1022
Db 1021 CCATATCTCCCTGACTTGTGTACAGCCACTGCTACCCATTTTGTGAGATGTTTCCCGATA 1080
QY 1023 TCCACTGGGCTCCCGCAGCCCGCAGTTTGGAGGCGAGGGGTGCCGAACAGCAGATCT 1082
Db 1081 TCCACTGGGCTCCCGCAGCCCGCAGTTTGGAGGCGAGGGGTGCCGAACAGCAGATCT 1140
QY 1083 CTGACCTGACACGAGGAGCCGAGTTGGAACCCGAGGAGAGAGACCGAGGTGCCAGGT 1142
Db 1141 CTGACCTGACACGAGGAGCCGAGTTGGAACCCGAGGAGAGAGACCGAGGTGCCAGGT 1200
QY 1143 ACCAATGGCATTCATGTCACCGGCGGTATGACTATCACTGAGCAACATCTCATCTAC 1202
Db 1201 ACCAATGGCATTCATGTCACCGGCGGTATGACTATCACTGAGCAACATCTCATCTAC 1260
QY 1203 AATGAGCACTATCTGGGGGGAGACACCGGGTCTGTGAGACCTCCAGTACCCCGAATCT 1262
Db 1261 AATGAGCACTATCTGGGGGGAGACACCGGGTCTGTGAGACCTCCAGTACCCCGAATCT 1320
QY 1263 CCATATCCCATTCGCGAAGAGGGGAGCCCTGCGGCTCCCGGGCTCTACACCCGACCG 1322
Db 1321 CCATATCCCATTCGCGAAGAGGGGAGCCCTGCGGCTCCCGGGCTCTCTACACCCGACCG 1380
QY 1323 GAAGATGAGCAAGCTTGGACCTTACCTAGCGGAGACAGACACTGTGTGCCACACCTTAAC 1382
Db 1381 GAAGATGAGCAAGCTTGGACCTTACCTAGCGGAGACAGACACTGTGTGCCACACCTTAAC 1440
QY 1383 AGGGGCCCAAGAAACCAATTTATCACCCATGACTGAGTGTGAGAAAGGAGAGAA 1442
Db 1441 AGGGGCCCAAGAAACCAATTTATCACCCATGACTGAGTGTGAGAAAGGAGAGAA 1500
QY 1443 AGGGGGGCAACAGGACCTTCTCCTTGAAGGTGCTGCCCCAGCTGGGATTCACAGGG 1502
Db 1501 AGGGGGGCAACAGGACCTTCTCCTTGAAGGTGCTGCCCCAGCTGGGATTCACAGGG 1560
QY 1503 GCTGTGATAGGGCGCGGGAGACGAGGCGCTTAAGGATTAAGGCTCAGACACCTTGAGA 1562
Db 1561 GCTGTGATAGGGCGCGGGAGAGAGGCGCTTAAGGATTAAGGCTCAGACACCTTGAGA 1620
QY 1563 GCAGGTGGGCACTGGCTGGGTACGGTCCCTCCAGAGACTCTTCCCTACTGTCCAGGAA 1622
Db 1621 GCAGGTGGGCACTGGCTGGGTACGGTCCCTCCAGAGACTCTTCCCTACTGTCCAGGAA 1680
QY 1623 ACCTGAGGCTCCCGGAGAACCCACCCACCCCTGGGGCTGTCAAGCTCAGGCAAGGAC 1682
Db 1681 ACCTGAGGCTCCCGGAGAACCCACCCACCCCTGGGGCTGTCAAGCTCAGGCAAGGAC 1740
QY 1683 AGGGACATGATACCAACTGTGCGCCACTACAGACGGCGGACCGGAGACGGGACCGAG 1742
Db 1741 AGGGACATGATACCAACTGTGCGCCACTACAGGACGGCGGACCGGAGACCGGACCGAG 1800
QY 1743 GAGGCGGCCACAGGTCACCTTGCAAGAGGTCACGCGGCCCTCTTAAAGATTCGTGTGC 1802
Db 1801 GAGGCGGCCACAGGTCACCTTGCAAGAGGTCACGCGGCCCTCTTAAAGATTCGTGTGC 1860
QY 1803 TCATCCCAAGGTCAGAGACCTTGGGGTTCACACTTCAAGTGAAGTGAAGTGAAGTGAAG 1862
Db 1861 TCATCCCAAGGTCAGAGACCTTGGGGTTCACACTTCAAGTGAAGTGAAGTGAAGTGAAG 1920

```

OY	51	GCCTGGAGGCGCGGCTGGCGGCTCCGGGCTCGGGGCGACATCGGCGCTGAGTCCG	110
Db	1	GCCTGGAGGCGCGGCTGGCGGCTCCGGGCTCGGGGCGACATCGGCGCTGAGTCCG	60
OY	111	TCCCAGGCTCTGGGCTCTGGGCAAGCCGCGCACCGCTGGCCAGAGAGCTCGGAGCTCTGC	170
Db	61	TCCCAGGCTCTGGGCTCTGGGCAAGCCGCGCACCGCTGGCCAGAGAGCTCGGAGCTCTGC	120
OY	171	CTTCTCTCCAGGCGCCGACGTTCTCTGGCGGCTGGGCGAGTGGCGGCATGCTCTGCT	230
Db	121	CTTCTCTCCAGGCGCCGACGTTCTCTGGCGGCTGGGCGAGTGGCGGCATGCTCTGCT	180
OY	231	TGGGCAACACTGTCGCCCGGCTGCGGCTGGGGGACCTTGATGCTGGGCTCTCGGAGCTC	290
Db	181	TGGGCAACACTGTCGCCCGGCTGCGGCTGGGGGACCTTGATGCTGGGCTCTCGGAGCTC	240
OY	291	CTGGACAGATCGACAGCGCCGACGCGGTGCTCCATATGCTGGAGAACAGACCTGACG	350
Db	241	CTGGACAGATCGACAGCGCCGACGCGGTGCTCCATATGCTGGAGAACAGACCTGACG	300
OY	351	GACCAGGAAAAAGAAATCTATGAGGCCCGAGACCGCATGTGTCGCCGCGCGCGCA	410
Db	301	GACCAGGAAAAAGAAATCTATGAGGCCCGAGACCGCATGTGTCGCCGCGCGCGCA	360
OY	411	GGACACTATGTTCTGAGTAAATGATGACCGCATCGGGACACAGTTTGTGCCATATGTC	470
Db	361	GGACACTATGTTCTGAGTAAATGATGACCGCATCGGGACACAGTTTGTGCCATATGTC	420
OY	471	GAGAAATCTCAACAGAGCACTGGAACTACAGCAATCGCAATCGGCACTGGCGGCGCTG	530
Db	421	GAGAAATCTCAACAGAGCACTGGAACTACAGCAATCGGCAATCGGCGGCGCTG	480
OY	531	GACCAAGTGAATGGGCTCTGAGAGATTTGGCCCTGCGACAGCAAAAGAGACCCAGTGC	590
Db	481	GACCAAGTGAATGGGCTCTGAGAGATTTGGCCCTGCGACAGCAAAAGAGACCCAGTGC	540
OY	591	CGCTGCCAGCGCGGAAATGTTCTGTGCTGGCTGGGCGCTGAGTGTACACACTGGAGCTA	650
Db	541	CGCTGCCAGCGCGGAAATGTTCTGTGCTGGCTGGGCGCTGAGTGTACACACTGGAGCTA	600
OY	651	CTTTCTGACTGCGCGCTGCGACTGTGAAGCGAGACTCAAAAGATGAGTTGGAGAGGTATAC	710
Db	601	CTTTCTGACTGCGCGCTGCGACTGTGAAGCGAGACTCAAAAGATGAGTTGGAGAGGTATAC	660
OY	711	AACCACTGCGTCCCTGTCAGAGGCACTTCAGAAATCTCTCTCCCGACGCGCGCG	770
Db	661	AACCACTGCGTCCCTGTCAGAGGCACTTCAGAAATCTCTCTCCCGACGCGCGCG	720
OY	771	TGCCAGCGCCACACAGGTGTGAAGCAAAAGGTGTGTGTGAAGGACGTCTCAGGCACTGCC	830
Db	721	TGCCAGCGCCACACAGGTGTGAAGCAAAAGGTGTGTGTGAAGGACGTCTCAGGCACTGCC	780
OY	831	CAGTCGACACAACTCTGCAAAAAATCCATTATGAGACCACTGCCCCGAGAGATGTCA-----	884
Db	781	CAGTCGACACAACTCTGCAAAAAATCCATTATGAGACCACTGCCCCGAGAGATGTCA	840
OY	885	-----	884
Db	841	ATGCTGATGCTGGCGCTTCTGCTGCCACTGCGCTTCTTCTGCTCCTTCCACCGCTTTC	900
OY	885	-----	902
Db	901	TCCTGCACTGTGGAAGACCACTCTCTCTCTGAGAAACTGGATCGCTGTCAGAGG	960
OY	903	CGTCCGAGGAGAGGAGACCACTCTCTGTAAGCTGGAAGCTGGAGCTCTCCGAGGCGCAT	962
Db	961	CGTCCGAGGAGAGGAGACCACTCTCTGTAAGCTGGAAGCTGGAGCTCTCCGAGGCGCAT	1020
OY	963	CCATATCTCCCTGACTGTGTGTAAGGACCACTGCAACCATTTCTGGAGATGTTTCCCACTA	1022
Db	1021	CCATATCTCCCTGACTGTGTGTAAGGACCACTGCAACCATTTCTGGAGATGTTTCCCACTA	1080
OY	1023	TCGCACTGGGCTCCCGCAGCCCAAGTTTGTGAAGGAGGGGTGCCCAACAGACAGAGTCT	1082

[illegible]

QY 351 GACCAAGAAAGAAATATATAGAGCCAGCAGCCGATCTGCTCCGCTGCCGCCA 410
 DB 301 GACCAAGAAAGAAATATATAGAGCCAGCAGCCGATCTGCTCCGCTGCCGCCA 360
 QY 411 GGCACCTATGTCTCAGCTAAATGTAGCCGATCCGGAGACAGATTGTGCGACATGTGCC 470
 DB 361 GGCACCTATGTCTCAGCTAAATGTAGCCGATCCGGAGACAGATTGTGCGACATGTGCC 420
 QY 471 GAGAAATCTTACAGAGAGAGCTGAACTACTGACATCTGACAGTGTGCCGCGCCCTGT 530
 DB 421 GAGAAATCTTACAGAGAGAGCTGAACTACTGACATCTGACAGTGTGCCGCGCCCTGT 480
 QY 531 GACCCAGATATGGGCTCGAGAGAGATTGCCCTCGACAGCAAGCAAGCAAGCAAGCAAG 590
 DB 481 GACCCAGATATGGGCTCGAGAGAGATTGCCCTCGACAGCAAGCAAGCAAGCAAGCAAG 540
 QY 591 GCGTCCAGAGCCGGGATGTCTGTCTGCTGGGCGCTCGAGTGTACACACTGCGAGCTA 650
 DB 541 GCGTCCAGAGCCGGGATGTCTGTCTGCTGGGCGCTCGAGTGTACACACTGCGAGCTA 600
 QY 651 CTTTCTGACGTGCGCCCTGGGCACTGAGCCGAGCTCAAGATGAGATTGGGAGGGTAA 710
 DB 601 CTTTCTGACGTGCGCCCTGGGCACTGAGCCGAGCTCAAGATGAGATTGGGAGGGTAA 660
 QY 711 AACCACTGGTCCCTCGAGAGGAGGAGCTTCCAGAAATACCTCTCCCGCAGCGCCGC 770
 DB 661 AACCACTGGTCCCTCGAGAGGAGGAGCTTCCAGAAATACCTCTCCCGCAGCGCCGC 720
 QY 771 TGCCAGCCCAACACAGAGTGTAGAACCAAGGTCTGGTGGAGGAGGAGCTCCAGGAGCTGC 830
 DB 721 TGCCAGCCCAACACAGAGTGTAGAACCAAGGTCTGGTGGAGGAGGAGCTCCAGGAGCTGC 780
 QY 831 CAGTCCGACACACCTGCAAAAATCCATTAGAGCCACTGCCCCAGAGATGTCA----- 884
 DB 781 CAGTCCGACACACCTGCAAAAATCCATTAGAGCCACTGCCCCAGAGATGTCAAGAAC 840
 QY 885 ----- 884
 DB 841 ATGCTGATGCTGGCCGTTTGTGCTGCCACTGGCCCTTCTTCTGCTCCAGCCGCTTC 900
 QY 885 -----GGATGGCTGCTCAAGAGG 902
 DB 901 TCTGCACTGTGAGAGGACCAACCTCTCTCTGAGAGAACTGGGATGCTGTCAAGAGG 960
 QY 903 CGTCCGAGAGGAGAGGAGCAATCTCTGATGAGGAGCTGGGAGCTCCGAAAGGCCAT 962
 DB 961 CGTCCGAGAGGAGAGGAGCAATCTCTGATGAGGAGCTGGGAGCTCCGAAAGGCCAT 1020
 QY 963 CCATACTTCCCTGACTGTGATAGAGCCACTGCTACCCATTTCTGAGAGATGTTTCCAGTA 1022
 DB 1021 CCATACTTCCCTGACTGTGATAGAGCCACTGCTACCCATTTCTGAGAGATGTTTCCAGTA 1080
 QY 1023 TCCACTGGGCTCCCGCAGCCAGTTTGTGAGGAGGAGGAGTGGCGCAACGAGAGCT 1082
 DB 1081 TCCACTGGGCTCCCGCAGCCAGTTTGTGAGGAGGAGGAGTGGCGCAACGAGAGCT 1140
 QY 1083 CTGAGACTGACAGGAGGAGCCGAGTTGGAACCCGGGAGAGCAGAGTGGGCCAGGT 1142
 DB 1141 CTGAGACTGACAGGAGGAGCCGAGTTGGAACCCGGGAGAGCAGAGTGGGCCAGGT 1200
 QY 1143 ACCAATGCAATTCATGTCACCGGAGTCTATGACTATCACTGGCAACATCTACATCTAC 1202
 DB 1201 ACCAATGCAATTCATGTCACCGGAGTCTATGACTATCACTGGCAACATCTACATCTAC 1260
 QY 1203 AATGAGCAGTACTGAGGAGGAGCAGCGGTCTCTGAGAGCTTCCACACTACCCCGAAGCT 1262
 DB 1261 AATGAGCAGTACTGAGGAGGAGCAGCGGTCTCTGAGAGCTTCCACACTACCCCGAAGCT 1320
 QY 1263 CCATACCCATTCGCGAAGAGGAGGAGCCTGGGCTCTCCCGGCTCTCTACACCCAGCAG 1322
 DB 1321 CCATACCCATTCGCGAAGAGGAGGAGCCTGGGCTCTCCCGGCTCTCTACACCCAGCAG 1380
 QY 1323 GAAGATGGCAAGGCTTGGCAGCTAGGAGAGACAGACACTGTGTGCCACACCTCTTAC 1382

DB 1381 GAAAGTGGCAAGGCTGTGGCACTAGCGAGACAGAGCACTGTGGTGCACACCTCTAAC 1440
 QY 1383 AGGGGCCCAAGAACCAATTATACCATATGACTAGAGTGTGAAAAAGAGAGAGA 1442
 DB 1441 AGGGGCCCAAGAACCAATTATATACCATATGACTAGAGTGTGAAAAAGAGAGAGA 1500
 QY 1443 AGGGGGGACACAGGGGACCTTCTCCCTTGAGAGCTGCCCCGACAGTGGATTACAGGG 1502
 DB 1501 AGGGGGGACACAGGGGACCTTCTCCCTTGAGAGCTGCCCCGACAGTGGATTACAGGG 1560
 QY 1503 GCTGAGTAGGGGCCCGGAGAGCAGAGCCCTTAAGGGATTAAAGGTGAGAACTGTAGA 1562
 DB 1561 GCTGAGTAGGGGCCCGGAGAGCAGAGCCCTTAAGGGATTAAAGGTGAGAACTGTAGA 1620
 QY 1563 GCAAGTGGGCACTGTGGGTAGCGTCCCTCCACAGAGCTCTCCACTAGCTGAGCAA 1622
 DB 1621 GCAAGTGGGCACTGTGGGTAGCGTCCCTCCACAGAGCTCTCCACTAGCTGAGCAA 1680
 QY 1623 ACCTGAGGCTCTCCCGGAGAGACCCACCCCTGGGCTGTGAGCTGAGGACGAG 1682
 DB 1681 ACCTGAGGCTCTCCCGGAGAGACCCACCCCTGGGCTGTGAGCTGAGGACGAG 1740
 QY 1683 AGGCAATGATATACCACTGCTGCCCACTACAGAGCGGCAAGGAGAGCAGCGAG 1742
 DB 1741 AGGCAATGATATACCACTGCTGCCCACTACAGAGCGGCAAGGAGAGCAGCGAG 1800
 QY 1743 GGAAGCCGACACAGGCTACCTGCAAGAGCTCAGGAGCCCTCTAAAGATTCTGTGTC 1802
 DB 1801 GGAAGCCGACACAGGCTACCTGCAAGAGCTCAGGAGCCCTCTAAAGATTCTGTGTC 1860
 QY 1803 TCAATCCCAAGCTTCAGAGACCTTTGGGTTTCAACATTTACGTGAGTGAAGTAC 1862
 DB 1861 TCAATCCCAAGCTTCAGAGACCTTTGGGTTTCAACATTTACGTGAGTGAAGTAC 1920
 QY 1863 CTCATGAGATGAATTTATAGAGAGAGCTCCCTCCCTCCCTGAGAGAGAG 1922
 DB 1921 CTCATGAGATGAATTTATAGAGAGAGAGCTCCCTCCCTCCCTGAGAGAGAG 1980
 QY 1923 AGGAGTCAATTAACAAGTGGGGTGGGTAGAGATCTAGAGTGGGAGAGATTGG 1982
 DB 1981 AGGAGTCAATTAACAAGTGGGGTGGGTAGAGATCTAGAGTGGGAGAGATTGG 2040

RESULT 4
 AX32212 2136 bp DNA linear PAT 09-JAN-2002
 LOCUS AX32212
 DEFINITION Sequence 2721 from Patent WO0194629.
 ACCESSION AX32212
 VERSION AX32212.1 GI:18122846
 KEYWORDS

SOURCE

human:

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.
 Cancer gene determination and therapeutic screening using signature
 gene sets
 Patent: WO 0194629-A 2721 13-DEC-2001;

JOURNAL
 Avalon Pharmaceuticals (US)
 FEATURES
 location/Qualifiers
 1..2136

BASE COUNT 446 a 706 c 608 g 376 t
 ORIGIN
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

Query Match 91.3%; Score 1809.2; DB 6; Length 2136;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

QY 688 AAGATGAAGTTGGGAAGGGTAACAACACACTGCTCCCTGCAAGGAGGAGGACTTCAGAGA 747
 Db 661 AAGATGAAGTTGGGAAGGGTAACAACACACTGCTCCCTGCAAGGAGGAGGACTTCAGAGA 720
 QY 748 ATACTCTCTCCCGAGGCGCGCGTCCGACCCACACAGAGTGTGAAGAACCAAGTCTGG 807
 Db 721 ATACTCTCTCCCGAGGCGCGCGTCCGACCCACACAGAGTGTGAAGAACCAAGTCTGG 780
 QY 808 TGAAGGACGCTCCAGGACCTGCGACAGTCCGACACAACTGTGAAAAATCATAGAGCCAC 867
 Db 781 TGAAGGACGCTCCAGGACCTGCGACAGTCCGACACAACTGTGAAAAATCATAGAGCCAC 840
 QY 868 TGCCCCCAGAGATGCA----- 884
 Db 841 TGCCCCCAGAGATGCTAGAGAACCAATGTGATGTGGCGCTTGTGCTGACACTGGCCTTCT 900
 QY 885 ----- 884
 Db 901 TTCTGCTCTTGGCAACCGCTCTTCTCTGATCTGGAAGAGCCACCTTCTCTGCAAGGA 960
 QY 885 -----GATCGCTGCTCAAGAGGGGTCGCGAGGGAGAGGAGCCCAATCTGTAGCTGGAA 939
 Db 961 AACTGGGATCCGCTGCTCAAGAGGGGTCGCGAGGGAGAGGAGCCCAATCTGTAGCTGGAA 1020
 QY 940 GCTGGGACCTCCGAAAGGCCCATCTACTTCCCTGACTGTGACAGCCACTGTACCCA 999
 Db 1021 GCTGGGACCTCCGAAAGGCCCATCTACTTCCCTGACTGTGACAGCCACTGTACCCA 1080
 QY 1000 TTTCTGGAGATGTTTCCCGAGTATCCTGAGTCTCCCGAGCCCGCACTTTTGGAGGAG 1059
 Db 1081 TTTCTGGAGATGTTTCCCGAGTATCCTGAGTCTCCCGAGCCCGCACTTTTGGAGGAG 1140
 QY 1060 GGGTCCCGCAACAGAGAGATGCTCTGAGACTGACAGGAGCCGAGTGGAAACCCGGGG 1119
 Db 1141 GGGTCCCGCAACAGAGAGATGCTCTGAGACTGACAGGAGCCGAGTGGAAACCCGGGG 1200
 QY 1120 AGCAGAGCCAGGATGGCCCGGATCAATGATGATGATGATGATGATGATGATGATGATGAT 1179
 Db 1201 AGCAGAGCCAGGATGGCCCGGATCAATGATGATGATGATGATGATGATGATGATGATGAT 1260
 QY 1180 TCACGTGGCAACATCTACATCTACATGATGATGATGATGATGATGATGATGATGATGAT 1239
 Db 1261 TCACGTGGCAACATCTACATCTACATGATGATGATGATGATGATGATGATGATGATGAT 1320
 QY 1240 AACTCCAGCTACCCCGGACCTCCATCCATCCATCCGAGAGGGGGAGCCCTGGCCCTC 1299
 Db 1321 AACTCCAGCTACCCCGGACCTCCATCCATCCATCCGAGAGGGGGAGCCCTGGCCCTC 1380
 QY 1300 CCGGGCTCTCTACACCCCGGACGAGAGATGGCAAGGCTTGGCACCCTAGCGAGACAGAC 1359
 Db 1381 CCGGGCTCTCTACACCCCGGACGAGAGATGGCAAGGCTTGGCACCCTAGCGAGACAGAC 1440
 QY 1360 ACTGTGTGACACACCTCTTAACAGGGGCCCAAGAGAACCAATTATACCATGACTGAC 1419
 Db 1441 ACTGTGTGACACACCTCTTAACAGGGGCCCAAGAGAACCAATTATACCATGACTGAC 1500
 QY 1420 TGAGTCTAGAAAAAGGAGGAAGAGGGGGGCAAAAGGGGACCTTCTCTGAGAGCTGCC 1479
 Db 1501 TGAGTCTAGAAAAAGGAGGAAGAGGGGGGCAAAAGGGGACCTTCTCTGAGAGCTGCC 1560
 QY 1480 CTGCCCCAGTGGGATTCACAGGGGGCTGAGTAGGGGCCCGGGGAACAGAGCCCTTAAGGA 1539
 Db 1561 CTGCCCCAGTGGGATTCACAGGGGGCTGAGTAGGGGCCCGGGGAACAGAGCCCTTAAGGA 1620
 QY 1540 TTAAGGCTCAAGACACCTCTGAGAGAGTGGGCACTGGCTGGGTACGGTCCCTCAGAC 1599
 Db 1621 TTAAGGCTCAAGACACCTCTGAGAGAGTGGGCACTGGCTGGGTACGGTCCCTCAGAC 1680
 QY 1600 GACTTCTCTACTGCTAGCAAAACCTGAGGGCTCCCGGAGAGCCACACCCCTCCGG 1659
 Db 1681 GACTTCTCTACTGCTAGCAAAACCTGAGGGCTCCCGGAGAGCCACACCCCTCCGG 1740
 QY 1660 GCTGCTCAGCTCAGGCAAGGAGGACATGATATCCAACTGCTGCCACTACAGCAG 1719

Db 1741 GCTGCTCAGCTCAGGCGACGAGAGGAGCAGATGATATACCAATGCTGCCACTACGAGCAG 1800
 QY 1720 CCGCAGCGGAGACCGGACCGAGGAGGCGCCACACGGTTCACCTGCAAGAGAGCTCAGGG 1779
 Db 1801 CCGCAGCGGAGACCGGACCGAGGAGGCGCCACACGGTTCACCTGCAAGAGAGCTCAGGG 1860
 QY 1780 CCCCTCTAAAGATTCGTGTGCTCATCCCAAGCTTCAGAGACCCTTTGGGGTTCCACA 1839
 Db 1861 CCCCTCTAAAGATTCGTGTGCTCATCCCAAGCTTCAGAGACCCTTTGGGGTTCCACA 1920
 QY 1840 CTTACAGTGGACTGAGTAGACCTCGCATGAAGATGAATATATAGAGAGAGCTCTTC 1899
 Db 1921 CTTACAGTGGACTGAGTAGACCTCGCATGAAGATGAATATATAGAGAGAGCTCTTC 1980
 QY 1900 CCTCCCTCCCTAGAGGAGAGGAGGAGGAGGATTAACAACTAAGGGGGTGGGTAGAGATTC 1959
 Db 1981 CCTCCCTCCCTAGAGGAGAGGAGGAGGAGGATTAACAACTAAGGGGGTGGGTAGAGATTC 2040
 QY 1960 CTAGGTATGGGAGAGATTTGG 1982
 Db 2041 CTAGGTATGGGAGAGATTTGG 2063

RESULT 3
 AX331906 2136 bp DNA linear PAT 09-JAN-2002
 LOCUS AX331906
 DEFINITION Sequence 2415 from Patent WO0194629.
 ACCESSION AX331906
 VERSION AX331906.1 GI:18122540
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
 AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets

JOURNAL Patent: WO 0194629-A 2415 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES location/Qualifiers
 source 1..2136

BASE COUNT 446 a 706 c 608 g 376 t
 ORIGIN

Query Match 91.3%; Score 1809.2; DB 6; Length 2136;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

QY 51 GCCCTGAGAGCCCGGCTGGCCGCTCCCGGCTGGGGTGCACATCGGCCCTGAGTCCG 110
 Db 1 GCCCTGAGAGCCCGGCTGGCCGCTCCCGGCTGGGGTGCACATCGGCCCTGAGTCCG 60
 QY 111 TCCAGAGCTTGGGCTCGGGGACGCGCGCCACCGCTCCCAAGAGAGTGGGCTCTGCG 170
 Db 61 TCCAGAGCTTGGGCTCGGGGACGCGCGCCACCGCTCCCAAGAGAGTGGGCTCTGCG 120
 QY 171 CTTCTCCAGGCGCCCGCATGTTGCTGGCGGCTGGCCAGTGGCGGCATCTCTGCTG 230
 Db 121 CTTCTCCAGGCGCCCGCATGTTGCTGGCGGCTGGCCAGTGGCGGCATCTCTGCTG 180
 QY 231 TGGGCACTCTGCCCCCGGCTGGGCGCTGTGGTGTGGGCTCTTGGGCTC 290
 Db 181 TGGGCACTCTGCCCCCGGCTGGGCGCTGTGGTGTGGGCTCTTGGGCTC 240
 QY 291 CTGGAGCATGGAGCCCGGAGGCGGTGCTCATATGCTGGGAGAACACAGCTGCAGG 350
 Db 241 CTGGAGCATGGAGCCCGGAGGCGGTGCTCATATGCTGGGAGAACACAGCTGCAGG 300

COMMENT	FEATURES	source
University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ams.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416) NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Depart of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).	Location/Qualifiers 1. 2148 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="HRC0478" /cell_type="primary human renal epithelial cells" /clone.lib="HRC" /note="cloning vector pME18FL3"	misc_feature 1. 2148 /note="highly similar to HUMTNRFP Homo sapiens tumor necrosis factor receptor 2 related protein mRNA"
BASE COUNT	462 a 705 c 606 g 375 t	
ORIGIN		
Query Match Best Local Similarity 94.7%; Pred. No. 0; Matches 1953; Conservative 0; Mismatches 2; Indels 108; Gaps 1;		
28 ACTTCCTGAGCTCCGGCATGGAGGAGCCCTGGAGGCGCGCTGCGCGCCGCGCTGGG 87		
1 ACTTCCTGAGCTCCGGCATGGAGGAGCCCTGGAGGCGCGCTGCGCGCCGCGCTGGG 60		
88 GTGGCAATGAGGCGCTGAGTCCGCTCCAGAGCTCTGGGCTGGGCGACCCGCGCCACCGCT 147		
61 GTGGCAATGAGGCGCTGAGTCCGCTCCAGAGCTCTGGGCTGGGCGACCCGCGCCACCGCT 120		
148 GCCCAAGAGAGTGGGGGCTCTGCTCCCTCCAGAGCCCGCCAGCTGGCGGCGCGCTGGGC 207		
121 GCCCAAGAGAGTGGGGGCTCTGCTCCCTCCAGAGCCCGCCAGCTGGCGGCGCGCTGGGC 180		
208 GAGTGGCGCCATGCTCTGCTGCTTGGGGCACCTTGCCCCGCGGCGCTGGGCGGCGCTC 267		
181 GAGTGGCGCCATGCTCTGCTGCTTGGGGCACCTTGCCCCGCGGCGCTGGGCGGCGCTC 240		
268 TGGTCTGGAGGCTCTTGGGGCTCCCGGAGAGATGAGACCCGAGGGGCTGCTCCATATG 327		
241 TGGTCTGGAGGCTCTTGGGGCTCCCGGAGAGATGAGACCCGAGGGGCTGCTCCATATG 300		
328 CGTGGAGAACCACTGTCAGGAGACCAAGAAAGGATATATAGACCCGAGCAGCGCA 387		
301 CGTGGAGAACCACTGTCAGGAGACCAAGAAAGGATATATAGACCCGAGCAGCGCA 360		
388 TCTGCTGCTCCCGCTGGCGCGGAGGACACTATGTCTAGCTAAATGTAGCCGCAATCCGG 447		
361 TCTGCTGCTCCCGCTGGCGCGGAGGACACTATGTCTAGCTAAATGTAGCCGCAATCCGG 420		
448 ACACAGTTTGTGCCATATGTCCGAGAAATTCCTCAACGAGACATGGAACTACTCGCA 507		
421 ACACAGTTTGTGCCATATGTCCGAGAAATTCCTCAACGAGACATGGAACTACTCGCA 480		
508 TCTGCACTGTGCCGCGCTGTACCCAGTGAAGGGCTCTGAGAGAGATTGCCCTGCA 567		
481 TCTGCACTGTGCCGCGCTGTACCCAGTGAAGGGCTCTGAGAGAGATTGCCCTGCA 540		
568 CAAGCAAAACGGAAGCCAGTGGCGGTGGCAAGCCGGGAATGTTGTGCTGCTGGGGCC 627		
541 CAAGCAAAACGGAAGCCAGTGGCGGTGGCAAGCCGGGAATGTTGTGCTGCTGGGGCC 600		
628 TCGAGTGTACACTGGAGACTATTTCTACTGCCCCGCTGGCACTGAAGCGGACACTCA 687		
601 TCGAGTGTACACTGGAGACTATTTCTACTGCCCCGCTGGCACTGAAGCGGACACTCA 660		

JOURNAL

Submitted (02-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH

CNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdickpaxil.stanford.edu
Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>.
Series: IRAL Plate: 37 Row: b Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505038.
Location/Qualifiers

1. .2161

/organism="Homo sapiens"
/db_xref="LocusID:4055"
/db_xref="taxon:9606"
/clone="MGC:22593 IMAGE:4703362"
/tissue_type="Lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
199. .1506
/codon_start=1
/product="Lymphotoxin beta receptor (TNFR superfamily, member 3)"
/protein_id="AAH262.1"
/db_xref="GI:20072213"

CDS

/translation="MLPWATSPAGLAWGELVGLLAASQAPVAPYASENQTCR
DDEKEYEDPHRICKSRCPGTIVSAKCSIRIDTVCATCAENSYNHNWLTQCLR
PCDPVAGLEIAFCTSRKRCQPCGFMCAALACTHCELLSDPGRTEALKDEY
GKNNHCVCQKAGHFOHTSSPARCOPHRCENMGIVAAQASDPTCKNLEPLP
PEMSGTMLAVILPLAFLLIATVSCITKSPISCRKIGSLIKRPOEGENPYAG
SMRPRAHYFPDLPVPLISGDSVSTGLPAALVAGVQDQSDLPDLREPLD
PGEQSVAVATNGIHTVGSMTITGNIYVNGVLTGAGTPGGLPATPPPYPIPEG
DPEPPLSTPHQEDGKAMHLETHCGAPSPNCRPNQFTHD"

BASE COUNT 473 a 708 c 606 g 374 t

ORIGIN

Query Match 92.8%; Score 1839.8; DB 9; Length 2161;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1959; Conservative 0; Mismatches 2; Indels 108; Gaps 1;

22 ACCGCCACTCTGAGCTCGCCATGGAGCCCTGGAGGCCCGCCCTCCGGC 81
Db 2 ACTGCCACTCTGAGCTCGCCATGGAGCCCTGGAGGCCCGCCCTCCGGC 61
82 CTTGGGGTGCATCGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 141
Db 62 CTTGGGGTGCATCGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 121
82 CTTGGGGTGCATCGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 141
Db 142 ACCGCTGCCAGAGCTGGGCTCTGCTCTCCAGGCGCCCGCAGTGTGGCCG 201
122 ACCGCTGCCAGAGCTGGGCTCTGCTCTCCAGGCGCCCGCAGTGTGGCCG 181
82 CTTGGGGTGCATCGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 141
Db 182 CTTGGGGTGCATCGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 121
262 GGCCTCTGAGTCTGGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 321
Db 242 GGCCTCTGAGTCTGGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 301

322 CATATGGCTCGAGAACACGACTGACAGGACACAGAAAGAAATATGAGCCCGC 381
Db 302 CATATGGCTCGAGAACACGACTGACAGGACACAGAAAGAAATATGAGCCCGC 361
382 ACCGATCTGCTGCTCCCGCTGCGCCAGGACCTATGTCAGTAAATATGAGCCG 441
Db 362 ACCGATCTGCTGCTCCCGCTGCGCCAGGACCTATGTCAGTAAATATGAGCCG 421
442 TCCGGGACACGTTTGGCCACATGTCGCGGAAATTCCTACAGAGACCTGGAAC 501
Db 422 TCCGGGACACGTTTGGCCACATGTCGCGGAAATTCCTACAGAGACCTGGAAC 481
502 TGACATCTGACAGCTGTCGCGCCCTGTGACAGGATGAGGCTCGAGAGATTCG 561
Db 482 TGACATCTGACAGCTGTCGCGCCCTGTGACAGGATGAGGCTCGAGAGATTCG 541
562 CTGACACAGCAAGCAAGAGACACAGTGCCTGTCGACAGCGGAAATGTCGTCCT 621
Db 542 CTGACACAGCAAGCAAGAGACACAGTGCCTGTCGACAGCGGAAATGTCGTCCT 601
622 GGGCTCTGAGTGTACACACTGAGGCTACTTGTGACTGCGCGCTGGACAGTAA 681
Db 602 GGGCTCTGAGTGTACACACTGAGGCTACTTGTGACTGCGCGCTGGACAGTAA 661
682 AGCTCAAGATGAAGTTGGGAAGGTTAACAACACTGCTCCCTGCAAGGACAGG 741
Db 662 AGCTCAAGATGAAGTTGGGAAGGTTAACAACACTGCTCCCTGCAAGGACAGG 721
742 TCCAGATTAATCTCTCTCCCGCCAGGCGCGCTGACAGCCCGACACAGTGTGA 801
Db 722 TCCAGATTAATCTCTCTCCCGCCAGGCGCGCTGACAGCCCGACACAGTGTGA 781
802 GTCTGTGTGAGAGAGCTCTCAGAGCTGCTCCAGTCCGACACAACTGCAAAAT 861
Db 782 GTCTGTGTGAGAGAGCTCTCAGAGCTGCTCCAGTCCGACACAACTGCAAAAT 841
862 AGCCACTGCCCCAGAGATGTCA----- 884
Db 842 AGCCACTGCCCCAGAGATGTCA----- 864
885 ----- 901
902 CTTTCTTCTGCTCTTCCACACGCTTCTCTGACATCTGGAAGAGCCACTTCTCT 961
Db 885 ----- 901
962 GCGAAGAACTGGGATCTGCTCTCTCAGAGGCTGCGAGGAGAGGAGCCCAATCT 1021
934 CTGAAAGCTGGAGGCTCTGCAAGGCTCCATCACTTCTGCTGATGTTGTTACAG 993
Db 1022 CTGAAAGCTGGAGGCTCTGCAAGGCTCCATCACTTCTGCTGATGTTGTTACAG 1081
994 TACCATTTCTGAGATGTTTCCCACTATCCACTGAGGCTCCCGACAGCCAGTT 1053
Db 1082 TACCATTTCTGAGATGTTTCCCACTATCCACTGAGGCTCCCGACAGCCAGTT 1141
1054 AGGAGAGGCTGCGCAACAGCAGAGTCTCTGACACTGACAGGAGGAGCCAGTT 1113
Db 1142 AGGAGAGGCTGCGCAACAGCAGAGTCTCTGACACTGACAGGAGGAGCCAGTT 1201
1114 CCGGGAGAGAGAGCAGGCTGCGCAGGCTGCAATGATGTCACCGGGGAGTCT 1173
Db 1202 CCGGGAGAGAGAGCAGGCTGCGCAGGCTGCAATGATGTCACCGGGGAGTCT 1261
1262 TGACTATCACTGCAACATCTACATCTACAAATGAGACAGTATGAGGAGGAG 1321
82 CTTGGGGTGCATCGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 141
Db 1234 CTTGGGGTGCATCGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 1293
1322 CTTGGGGTGCATCGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 1381
82 CTTGGGGTGCATCGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 141
Db 1294 GGCCTCTGAGTCTGGGCTCTGAGTCCGTCAGGCTCTGGGCTGGAGCGCGCC 1353

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:04:15 ; Search time 3695.99 seconds
(without alignments)
15606.600 Million cell updates/sec

Title: US-09-917-372-2

Perfect score: 1982

Sequence: 1 gccccgcgcgcagctcgtc.....ggtatggggaagagtttgg 1982

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_to:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlg_hum:*
40: em_hlg_mus:*
41: em_hlg_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1839.8	92.8	2161	9	BC026262
2	1833.8	92.5	2148	6	AK027080
3	1809.2	91.3	2136	6	AX331906
4	1809.2	91.3	2136	6	AX332212
5	1809.2	91.3	2136	6	AX409488
6	1809.2	91.3	2136	6	HMTNFRRP
7	1732.2	87.4	2091	9	AK095418
8	841.4	42.5	140026	9	AC005840
9	765.8	38.6	2076	10	MMU25173
10	673.8	34.0	1614	10	MUSLYMPHOB
11	444.6	22.4	527	6	AX381072
12	296.8	15.0	400	11	G11431
13	281.8	14.2	178228	2	AC128082
14	264	13.3	187998	2	AC125809
15	198	10.0	1605	9	HMTNFRNEC
16	178.8	9.0	250	11	G11535
17	151.2	7.6	187998	2	AC125909
18	149.6	7.5	2048	10	MMU30798
19	99	5.0	323	9	HS2789R
20	97.6	4.9	193	9	HS2789R
21	96.8	4.9	178228	2	AC128082
22	73.4	3.7	125020	9	AF429315
23	72	3.6	1557	6	AX054825
24	72	3.6	1557	6	I36197
25	72	3.6	1641	6	I36196
26	72	3.6	2224	6	AR152033
27	72	3.6	2253	6	A78517
28	72	3.6	2394	9	HMTNFRRII
29	72	3.6	3380	11	G26865
30	72	3.6	3683	6	AX333705
31	72	3.6	3683	6	AX348016
32	72	3.6	3683	6	AX348018
33	72	3.6	3683	6	AX348020
34	72	3.6	3683	9	HMTNFR
35	69.4	3.5	125020	9	AF429315
36	62.6	3.2	691	6	I36350
37	62.6	3.2	705	6	AR174443
38	62.4	3.1	579	6	AR168180
39	62.4	3.1	591	6	AR168183
40	62.4	3.1	852	9	AF373877
41	62.4	3.1	852	9	AF373878
42	62.4	3.1	1596	6	AR168182
43	62.4	3.1	1621	9	BC002794
44	62.4	3.1	1724	6	AR169913
45	62.4	3.1	1724	6	AR172617

ALIGNMENTS

RESULT 1
LOCUS BC026262 2161 bp mRNA linear PRI 08-APR-2002
DEFINITION Homo sapiens, lymphotoxin beta receptor (TNFR superfamily, member 3), clone MGC:22593 IMAGE:4703362, mRNA, complete cds.
ACCESSION BC026262
VERSION BC026262.1 GI:20072212
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2161)
AUTHORS Strausberg, R.
TITLE Direct Submission

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:04:15 ; Search time 792.53 Seconds

(without alignments)
15606.600 Million cell updates/sec

Title: US-09-917-372-2_COPY_853_1277

Perfect score: 425
Sequence: 1 atccatcagaccactgccc.....aacctccatcaccattccc 425

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hlg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pal: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hlg_hum: *
31: em_hlg_inv: *
32: em_hlg_other: *
33: em_hlg_mus: *
34: em_hlg_pln: *
35: em_hlg_rod: *
36: em_hlg_mam: *
37: em_hlg_vrt: *
38: em_sy: *
39: em_hlgo_hum: *
40: em_hlgo_mus: *
41: em_hlgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	393	92.5	2091	9 AK095418	AK095418 Homo sapi
2	393	92.5	2136	6 AX331906	AX331906 Sequence
3	393	92.5	2136	6 AX332212	AX332212 Sequence
4	393	92.5	2136	6 AX409488	AX409488 Sequence
5	393	92.5	2136	6 HUMTNRFP	L04270 Homo sapien
6	393	92.5	2148	9 AK027080	AK027080 Homo sapi
7	393	92.5	2161	9 BC026262	BC026262 Homo sapi
8	235.4	55.4	140026	9 AC005840	AC005840 Homo sapi
9	219.2	51.6	1614	10 MUSLYMPHOB	L38423 Mus musculu
10	219.2	51.6	2076	10 MMU29173	U29173 Mus musculu
11	117.6	27.7	178228	2 AC128082	AC128082 Rattus no
12	82.4	19.4	187998	2 AC125909	AC125909 Rattus no
13	40	9.4	125020	9 AF429315	AF429315 Homo sapi
14	38.4	9.0	125020	9 AF429315	AF429315 Homo sapi
15	37.8	8.9	6349	5 GGTNY	AF429315 Homo sapi
16	37.8	8.9	61851	9 AC022240	X99062 G.gallus mr
17	37.8	8.9	159950	2 AC016413	AC022240 Homo sapi
18	37.8	8.9	179357	9 AC074191	AC016413 Homo sapi
19	37.4	8.8	66808	1 SAU421825	AC074191 Homo sapi
20	36.8	8.7	232957	2 AC106606	AJ421825 Stigmatel
21	36.6	8.6	151437	2 AC118809	AC106606 Rattus no
22	36	8.5	170302	2 HST402N21	AC118809 Rattus no
23	36	8.5	187336	2 AP005427	AL049553 Human DNA
24	35.6	8.4	62496	2 AC090664	AP005427 Oryza sat
25	35.6	8.4	97120	9 AP004147	AC090664 Homo sapi
26	35.6	8.4	102173	9 AP000796	AP004147 Homo sapi
27	35.6	8.4	221357	2 AP004246	AP000796 Homo sapi
28	35.4	8.3	60660	9 AC005903	AP004246 Homo sapi
29	35.4	8.3	64784	2 AC036134	AC005903 Homo sapi
30	35.4	8.3	135622	9 AC093128	AC036134 Homo sapi
31	35.4	8.3	207541	2 AC121252	AC093128 Papio cyn
32	35.4	8.3	218687	2 AC079501	AC121252 Homo sapi
33	35.4	8.3	218753	2 AC079515	AC079501 Mus muscu
34	35.2	8.3	333	6 AX002166	AC079515 Mus muscu
35	35.2	8.3	333	9 HSY12721	AX002166 Sequence
36	35.2	8.3	36631	9 HUMCOL7A1X	Y12721 H.sapiens m
37	35.2	8.3	164407	2 AC097229	L23982 Homo sapien
38	35.2	8.3	232784	2 AC113522	AC097229 Sus scrof
39	35	8.2	1336	3 AF350267	AC113522 Mus muscu
40	35	8.2	127178	9 AC005160	AF350267 Argiope t
41	35	8.2	146170	2 AC079224	AC005160 Homo sapi
42	35	8.2	153570	2 AC027307	AC079224 Homo sapi
43	35	8.2	230372	2 AC073693	AC027307 Homo sapi
44	34.8	8.2	936	9 AF143881	AC073693 Mus muscu
45	34.8	8.2	6150	10 MMU76618	AF143881 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AK095418 2091 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens CDNA FLJ38099 fls. clone D3OST1000238, highly similar
to LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
ACCESSION AK095418
VERSION AK095418.1 GI:21754669
KEYWORDS Oligo capping; fls (full insert sequence).
SOURCE Homo sapiens CD34+ Cells CDNA to mRNA, clone_1lb:D3OST1
clone:D3OST1000238.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,

Arta, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashta, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Matsumura, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.

NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 2091)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB) (supported by Japan Construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
source
1. 2091
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="D3OST1000238"
/cell_type="CD34+ Cells"
/clone_id="D3OST1"
/note="cloning vector: PME18SFL3-mRNA from CD34+ cells after 3-days ODF induction. Primary culture, CD34+ cells"

BASE COUNT
440 a 691 c 594 g 366 t

ORIGIN
Query Match 92.5%; Score 393; DB 9; Length 2091;
Best Local Similarity 100.0%; Pred. No. 2.1e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 GGATGCTGCTCAAGAGGCGTCCGAGAGAGAGGAGCAATCTGTACTGGAGGCTG 92
Db 930 GGATGCTGCTCAAGAGGCGTCCGAGAGAGAGGAGCAATCTGTACTGGAGGCTG 989
QY 93 GAGCCTCGAAGAGCCATCATCTACTTCCCTGACTTGGTGTACAGCCACTGCTACCAATTTCT 152
Db 990 GAGCCTCGAAGAGCCATCATCTACTTCCCTGACTTGGTGTACAGCCACTGCTACCAATTTCT 1049
QY 153 GGAGATGTTTCCCAATATCATGCTGGGCTCCCGACAGCCCAAGTTTGGAGGAGGGGTG 212
Db 1050 GGAGATGTTTCCCAATATCATGCTGGGCTCCCGACAGCCCAAGTTTGGAGGAGGGGTG 1109
QY 213 CGGCAACAGCAGAGTCTGACCTGACCTGACAGGAGCCAGTGTGAAACCGGGGAGCAG 272
Db 1110 CGGCAACAGCAGAGTCTGACCTGACCTGACAGGAGCCAGTGTGAAACCGGGGAGCAG 1169
QY 273 AGCCAGGTGGCCACGATGACCAATGTCATGTCACCGGGGTCTATGATCATCTACT 332
Db 1170 AGCCAGGTGGCCACGATGACCAATGTCATGTCACCGGGGTCTATGATCATCTACT 1229
QY 333 GGCACATCTACTATCTACATGAGCAGGATGAGTGGGGGAGCCACCGGGGTCTCTGAGACCTC 392
Db 1230 GGCACATCTACTATCTACATGAGCAGGATGAGTGGGGGAGCCACCGGGGTCTCTGAGACCTC 1289
QY 393 CCAGCTACCCCGAAGCTCCATATCCCATTTCCC 425
Db 1290 CCAGCTACCCCGAAGCTCCATATCCCATTTCCC 1322

RESULT 2
AX331906 2136 bp DNA linear PAT 09-JAN-2002
LOCUS AX331906
DEFINITION Sequence 2415 from Patent WO0194629.
ACCESSION AX331906

VERSION AX331906.1 GI:18122540
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrikan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature gene sets
Patent: WO 0194629-A 2415 13-DEC-2001;
Avalon Pharmaceuticals (US)
location/Qualifiers
1. 2136
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
BASE COUNT 446 a 706 c 608 g 376 t

ORIGIN
Query Match 92.5%; Score 393; DB 6; Length 2136;
Best Local Similarity 100.0%; Pred. No. 2.1e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 GGATGCTGCTCAAGAGGCGTCCGAGAGAGAGGAGCAATCTGTACTGGAGGCTG 92
Db 943 GGATGCTGCTCAAGAGGCGTCCGAGAGAGAGGAGCAATCTGTACTGGAGGCTG 1002
QY 93 GAGCCTCGAAGAGCCATCATCTACTTCCCTGACTTGGTGTACAGCCACTGCTACCAATTTCT 152
Db 1003 GAGCCTCGAAGAGCCATCATCTACTTCCCTGACTTGGTGTACAGCCACTGCTACCAATTTCT 1062
QY 153 GGAGATGTTTCCCAATATCATGCTGGGCTCCCGACAGCCCAAGTTTGGAGGAGGGGTG 212
Db 1063 GGAGATGTTTCCCAATATCATGCTGGGCTCCCGACAGCCCAAGTTTGGAGGAGGGGTG 1122
QY 213 CGGCAACAGCAGAGTCTGCTGACCTGACAGGAGGAGCCAGTGTGAAACCGGGGAGCAG 272
Db 1123 CGGCAACAGCAGAGTCTGCTGACCTGACAGGAGGAGCCAGTGTGAAACCGGGGAGCAG 1182
QY 273 AGCCAGGTGGCCACGATGACCAATGTCATGTCACCGGGGTCTATGATCATCTACT 332
Db 1183 AGCCAGGTGGCCACGATGACCAATGTCATGTCACCGGGGTCTATGATCATCTACT 1242
QY 333 GGCACATCTACTATCTACATGAGCAGGATGAGTGGGGGAGCAGCGGGTCTCTGAGACCTC 392
Db 1243 GGCACATCTACTATCTACATGAGCAGGATGAGTGGGGGAGCAGCGGGTCTCTGAGACCTC 1302
QY 393 CCAGCTACCCCGAAGCTCCATATCCCATTTCCC 425
Db 1303 CCAGCTACCCCGAAGCTCCATATCCCATTTCCC 1335

RESULT 3
AX332212 2136 bp DNA linear PAT 09-JAN-2002
LOCUS AX332212
DEFINITION Sequence 2721 from Patent WO0194629.
ACCESSION AX332212
VERSION AX332212.1 GI:18122846
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrikan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature gene sets
Patent: WO 0194629-A 2721 13-DEC-2001;
Avalon Pharmaceuticals (US)
location/Qualifiers
1. 2136

FEATURES
source

ORIGIN	BASE COUNT	462 a	705 c	606 g	375 t
Query Match	92.5%	Score 393:	DB 9:	Length 2148:	
Best Local Similarity	100.0%	Pred. No. 2, 1e-92:			
Matches 393:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:	
QY	33	GGATCGCTGCTCAAGAGAGCGCTCCGACGAGGAGAGAGGAGGCCAATCTCTGTAGCTGGAACCTGG	92		
DB	966	GGATCGCTGCTCAAGAGAGCGCTCCGACGAGGAGAGGAGGCCAATCTCTGTAGCTGGAACCTGG	1025		
QY	93	GAGCTCCGAAGGCGCCATCTCACTTCCCTGACTTGTATAGACACATGCTACCATTTTCT	152		
DB	1026	GAGCTCCGAAGGCGCCATCTCACTTCCCTGACTTGTATAGACACATGCTACCATTTTCT	1085		
QY	153	GGAGATGTTTCCCAAGATATCACTGGGCTCCCGACGCCCAAGTTTGAAGGAGGGGTG	212		
DB	1086	GGAGATGTTTCCCAAGATATCACTGGGCTCCCGACGCCCAAGTTTGAAGGAGGGGTG	1145		
QY	213	CCGCAACAGCAGAGTCTCTGTGACCTGACCCAGGAGGCGGACAGTTGGAACCCGGGAGACAG	272		
DB	1146	CCGCAACAGCAGAGTCTCTGTGACCTGACCCAGGAGGCGGACAGTTGGAACCCGGGAGACAG	1205		
QY	273	AGCCAGTGGGCCACGATGACCAATGGCAATGGCAATTCATGTCAACGGCGGGTCTATGACTATCACT	332		
DB	1206	AGCCAGTGGGCCACGATGACCAATGGCAATGGCAATTCATGTCAACGGCGGGTCTATGACTATCACT	1265		
QY	333	GGCAACATCTACATCTACAAATGGACCACTACTTGGGGGAGACACAGGGGCTCGAGACCTC	392		
DB	1266	GGCAACATCTACATCTACAAATGGACCACTACTTGGGGGAGACACAGGGGCTCGAGACCTC	1325		
QY	393	CCAGCTACCCCGGACGCTCCATACCCCAATCTCC	425		
DB	1326	CCAGCTACCCCGGACGCTCCATACCCCAATCTCC	1358		
RESULT 7					
LOCUS	BC026262	2161 bp	mRNA	linear	PRI 08-APR-2002
DEFINITION					Homo sapiens, lymphotoxin beta receptor (TNFR superfamily, member
ACCESSION	BC026262				3) clone MGC:125593 IMAGE:4703362, mRNA, complete cds.
VERSION	BC026262.1				GI:20072212
KEYWORDS					MGC.
SOURCE					Homo sapiens.
ORGANISM					Homo sapiens
REFERENCE					Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS					Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE					1 (bases 1 to 2161)
JOURNAL					Strausberg, R.
					Direct Submissions
					Submitted (02-APR-2002) National Institutes of Health, Mammalian
					Gene Collection (MGC), Cancer Genomics Office, National Cancer
					Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
					USA
REMARK					NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT					Contact: MGC help desk
					Email: cgabs-r@mail.nih.gov
					Tissue Procurement: CLONTECH
					cDNA Library Preparation: CLONTECH Laboratories, Inc.
					cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
					DNA Sequencing by: Sequencing Group at the Stanford Human
					Center, Stanford University School of Medicine, Stanford, CA 94305
					Web site: http://www.shgc.stanford.edu
					Contact: (Dickson, Mark) mdc@paxil.stanford.edu
					Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
					R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: <http://image.llnl.gov>
 Series: IRAL Plate: 37 Row: b Column: 13
 This clone was selected for full length sequencing because it
 passes the following selection criteria: matched mRNA 91: 4505038.

FEATURES
SOURCE

Location/Qualifiers
1. .2161
/organism="Homo sapiens"
/db_xref="locusID:4055"
/db_xref="taxon:9606"
/clone="MGC:22593 IMAGE:4703362"
/tissue="type="Lung"
/clone_id="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
199. .1506
/product="lymphotoxin beta receptor (TNFR superfamily, member 3)"
/codon_start=1
/protein_id="AAH26262.1"
/db_xref="GI:20072213"
/translation="MLFWATSAPLANGPLVLGFLILASQPAVPPVASENCTCR
DQEKVEYEPHRIICSSRCPTGVYSAKCSIRDTVCATCAENSYEHMNYLTICLCR
PCDPMVGLIEIAPCTSKRTQRCOPMFCAMALECTHCELSQCPGTAEIKDEV
GKNNHCYVCKAGHONTSSPARGCOPHTRCENGLVEARGLAOSDTCKNPLEPLP
PMSGCTMLAVLPLAFILLATVYFSCIRKSHSLCKRLSLKRRPQSGPNVAG
SWEPKAPHPFDLQPLPLISGVSPVSTGLPAPVLEAGVPQOQSLDTREOLE
PQESQVAHNGIHTVGSMTITGNIYINGVPGVGGPGGDDLPATPEPPYPIPEG
DPGPGISTPHQEDGKMWHLAETHECATPSNRPNQFIYHD"

BASE COUNT 473 a 708 c 606 g 374 t
ORIGIN

Query Match 92.5%; Score 393; DB 9; Length 2161;
Best Local Similarity 100.0%; Pred. No. 21e-92;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 GGATGCGTCTCAAGAGGGGTCCGAGAGGAGGAGCCCAATCCTACTGGAAGCTGG 92
|||||
Db 973 GGATGCTGCTCAAGAGGGGTCCGAGAGGAGGAGCCCAATCCTACTGGAAGCTGG 1032
QY 93 GAGCCTCCGAGAGGCCCATCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCAATTCT 152
|||||
Db 1033 GAGCCTCCGAGAGGCCCATCATCTACTTCCCTGACTTGGTACAGCCACTGCTACCAATTCT 1092
QY 153 GGAGATGTTTCCCGAGATATCCACTGGGCTCCCGCAGCCCAAGTTTGGAGCGAGGGGTG 212
|||||
Db 1093 GGAGATGTTTCCCGAGATATCCACTGGGCTCCCGCAGCCCAAGTTTGGAGCGAGGGGTG 1152
QY 213 CCGCAACAGACAGAGTCTTGACCTGACCGAGGAGCGGAGTGGATGGAACCCGGGGAGCAG 272
|||||
Db 1153 CCGCAACAGACAGAGTCTTGACCTGACCGAGGAGCGGAGTGGATGGAACCCGGGGAGCAG 1212
QY 273 AGCCAGGTGGCCACAGGTACCAATGGCATGTCTACCGGGGGTCTATGATATCACT 332
|||||
Db 1213 AGCCAGGTGGCCACAGGTACCAATGGCATGTCTACCGGGGGTCTATGATATCACT 1272
QY 333 GGCAACATCTACATCTACCAATGAGCAGTACTGGGGGAGCACCAGGGTCTCTGAGACCTTC 392
|||||
Db 1273 GGCAACATCTACATCTACCAATGAGCAGTACTGGGGGAGCACCAGGGTCTCTGAGACCTTC 1332
QY 393 CCAGCTACCCCGAGACCTCATACCCCATTTCCC 425
|||||
Db 1333 CCAGCTACCCCGAGACCTCATACCCCATTTCCC 1365

RESULT 8
AC005840/c 140026 bp DNA linear PRI 20-OCT-2000
LOCUS Homo sapiens complete sequence of a PAC clone RPI-102E24 containing
DEFINITION SYBL, CD27, and SCNNIA genes.
ACCESSION AC005840
VERSION AC005840.2 GI:10938025
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 140026)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Montgomery, K.T., Lau, S.T. and Kucherlapati, R.
High Throughput Sequencing of Human Chromosome 12
Unpublished
2 (bases 1 to 140026)
Montgomery, K.T., Lau, S.T. and Kucherlapati, R.
Direct Submission
Submitted (22-OCT-1998) Department of Molecular Biology, Albert
Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY
10461, USA
3 (bases 1 to 140026)
Montgomery, K.T., Lau, S.T. and Kucherlapati, R.
Direct Submission
Submitted (20-OCT-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY
10461, USA
On Oct 20, 2000 this sequence version replaced gi:3779004.
-----Genome Center:
Albert Einstein College of Medicine
Code: AECOM
Web site: <http://sequence.aecom.yu.edu/chr12/>
Contact: jhan@sequence.aecom.yu.edu

CLONE LENGTH: This sequence represents the entire insert of this
clone unless otherwise noted. If there are overlapping clones, the
overlaps are noted in the beginning and end of the Features
listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550).
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for human sequences.

Genes and Regions of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST
and cDNA sequences in Unigene. Genes demonstrate at least two exons
flanked by consensus splice sites that maintain sequence continuity
across the splice junctions. Sequences that are not identical
matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double
stranded sequence for all regions. All sequence is completed to a
standard of coverage with a minimum of 3 reads with no ambiguities.
If the sequence coverage for a region does not meet this standard,
it is indicated in the annotation as low coverage. Low coverage
linkages are verified by PCR product size verification or
verification of forward and reverse reads from clones which span
the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated average error rate is less than 1 per 10,000
bases using the Consed quality parameters. Regions that do not
meet this requirement are annotated as low quality.

-----Summary Statistics

Center project name: RPI-102E24
Sequencing vector: M13
Chemistry: Dye-terminator Big Dye; 100%
Assembly program: Phrap version 0.990319
Contig length: 140026
Fraction of Phrap value < 40: 0.0426
Error Rate in Consed: 0.56 per 10,000 bases
Number of N's in consensus: 1

----- Distribution of Quality < 40 Bases:

10001	*	*	*	*
9001	*	*	*	*
8001	*	*	*	*
7001	*	*	*	*
#	*	*	*	*
6001	*	*	*	*
5001	*	*	*	*
4001	*	*	*	*
3001	*	*	*	*
2001	*	*	*	*

ORGANISM

Mus musculus

REFERENCE 1 (bases 1 to 1614)
 AUTHORS Nakamura,T., Tashiro,K., Nazarea,M., Nakano,T., Sasayama,S. and Honjo,T.
 TITLE The murine lymphotoxin-beta receptor cDNA: isolation by the signal sequence trap and cDNA mapping

JOURNAL Genomics 30 (2), 312-319 (1995)
 MEDLINE 96163885
 PUBMED 8586432

FEATURES

1..1614
 location/Qualifiers

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_line="St-2"

/note="putative"

/product="lymphotoxin-beta receptor"

/protein_id="AAB00846.1"

/translation="MRLPASSPCGLAMGPILLGLSLVASOPOLVPPYRIENOTCM
 DQKEYEPMHDDCCSRCPGEFVAVCSRODPVCKPNSYNEMHNHSTCOLR
 PCDVLGFEFVAPCTSDRKACRCOPGMSCYLDNECVHCEERLVLCQPTAEVTD
 EIMDTDNCVPCRGHONTSSPRARCPHRTCEIOGLVEAPSTYSDTCKNPPR
 GAMLLILILSLVFLFTVTLACAMNHPSLCRKLGLTLKRHEGESPPCPAPRAD
 PHFDPLAEPLPMSGDLSPSPAGPTAPSLIEVVLQOQSPVQARELEAEPEGHGOVA
 HGANGIHVGTGSVTVGNITLYNGPVLVGTGTPDPPAPPEPYPPTPEGAPGSELS
 TPYQEDGKAMHLEETELGCDL"

CDS

BASE COUNT 316 a 537 c 455 g 306 t
 ORIGIN

Query Match 51.6%; Score 219.2; DB 10; Length 1614;
 Best Local Similarity 79.3%; Pred. No. 7.1e-47;
 Matches 260; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 97 CTCGAGAGCCCATCTACTCTGACTGTGACGACCATCTGACCATTTCTGGAG 156
 DB 1000 CTCGAGAGCCCATCTACTCTGACTGTGACGACCATCTGACCATTTCTGGAG 1059
 QY 157 ATGTTTCCCACTATCCATGCTGCGGCTCCCGCAGCCCAAGTTTGGAGGAGGCGTGC 216
 DB 1060 ACTGTGCTCCCATCTGCTGCTGAGACCCCAAGGCGCCCTCTGAGGAGAGTGTCTAC 1119
 QY 217 AACAGCAGAGTCTCTGACCTGACGAGGAGCCGACGTTGGAACCCGGGAGCAGAGCC 276
 DB 1120 AACAGCAGAGTCTCTGCTGCTGAGACCCCAAGGCGCCCTCTGAGGAGAGTGTCTAC 1179
 QY 277 AGGTGGCCACGCTGACATGTCATGTCACCGGCGGCTCTGACTGTGACCTGAGCA 336
 DB 1180 AGGTGGCCACGCTGCTGCTGAGACCCCAAGGCGCCCTCTGACTGTGACCTGAGCA 1239
 QY 337 ACATCTACATCTACATGAGCAGTACTGAGGAGGAGCACCAGGCTCTGAGAGCTCCAG 396
 DB 1240 ACATCTACATCTACATGAGGAGTACTGAGGAGGAGCACCAGGCTCTGAGAGCTCCAG 1299
 QY 397 CTACCCCGAAGCTCTCATACCCCATTC 424
 DB 1300 CTACCCCGAAGCTCTCATACCCCATTC 1327

RESULT 10

MMU29173

LOCUS MMU29173 2076 bp mRNA linear ROD 28-JUN-1995
 DEFINITION Mus musculus lymphotoxin-beta receptor mRNA, complete cds.
 ACCESSION U29173
 VERSION U29173.1 GI:881620
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2076)
 AUTHORS Force,W.R., Williams-Abbott,L., Browning,J., Hession,C., Tizard,R. and Ware,C.F.

TITLE Cloning and Characterization of the Mouse Lymphotoxin-beta Receptor
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2076)
 AUTHORS Force,W.R.

TITLE Direct Submission
 JOURNAL Submitted (15-JUN-1995) Walker R. Force, Biomedical Sciences, U.C. Riverside, Riverside, CA 92521-0121, USA

FEATURES

1..2076
 location/Qualifiers

/organism="Mus musculus"

/strain="CVB"

/db_xref="taxon:10090"

/sex="female"

/tissue-type="lung"

/note="transmembrane protein"

/codon_start=1

/product="lymphotoxin-beta receptor"

/protein_id="AAB68964.1"

/db_xref="GI:881621"

/translation="MRLPASSPCGLAMGPILLGLSLVASOPOLVPPYRIENOTCM
 DQKEYEPMHDDCCSRCPGEFVAVCSRODPVCKPNSYNEMHNHSTCOLR
 PCDVLGFEFVAPCTSDRKACRCOPGMSCYLDNECVHCEERLVLCQPTAEVTD
 EIMDTDNCVPCRGHONTSSPRARCPHRTCEIOGLVEAPSTYSDTCKNPPR
 GAMLLILILSLVFLFTVTLACAMNHPSLCRKLGLTLKRHEGESPPCPAPRAD
 PHFDPLAEPLPMSGDLSPSPAGPTAPSLIEVVLQOQSPVQARELEAEPEGHGOVA
 HGANGIHVGTGSVTVGNITLYNGPVLVGTGTPDPPAPPEPYPPTPEGAPGSELS
 TPYQEDGKAMHLEETELGCDL"

CDS

BASE COUNT 436 a 652 c 597 g 391 t
 ORIGIN

Query Match 51.6%; Score 219.2; DB 10; Length 2076;
 Best Local Similarity 79.3%; Pred. No. 7e-47;
 Matches 260; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 97 CTCGAGAGCCCATCTACTCTGACTGTGACGACCATCTGACCATTTCTGGAG 156
 DB 984 CTCGAGAGCCCATCTACTCTGACTGTGACGACCATCTGACCATTTCTGGAG 1043
 QY 157 ATGTTTCCCACTATCCATGCTGCGGCTCCCGCAGCCCAAGTTTGGAGGAGGCGTGC 216
 DB 1044 ACTGTGCTCCCATCTGCTGCTGAGACCCCAAGGCGCCCTCTGAGGAGAGTGTCTAC 1103
 QY 217 AACAGCAGAGTCTCTGACCTGACGAGGAGCCGACGTTGGAACCCGGGAGCAGAGCC 276
 DB 1104 AACAGCAGAGTCTCTGCTGCTGAGACCCCAAGGCGCCCTCTGAGGAGAGTGTCTAC 1163
 QY 277 AGGTGGCCACGCTGACATGTCATGTCACCGGCGGCTCTGACTGTGACCTGAGCA 336
 DB 1164 AGGTGGCCACGCTGCTGCTGAGACCCCAAGGCGCCCTCTGACTGTGACCTGAGCA 1223
 QY 337 ACATCTACATCTACATGAGCAGTACTGAGGAGGAGCACCAGGCTCTGAGAGCTCCAG 396
 DB 1224 ACATCTACATCTACATGAGGAGTACTGAGGAGGAGCACCAGGCTCTGAGAGCTCCAG 1283
 QY 397 CTACCCCGAAGCTCTCATACCCCATTC 424
 DB 1284 CTACCCCGAAGCTCTCATACCCCATTC 1311

RESULT 11

AC128082/C

LOCUS AC128082 178228 bp DNA linear HTG 19-JUL-2002
 DEFINITION Rattus norvegicus clone CH230-362C16, *** SEQUENCING IN PROGRESS
 ACCESSION AC128082
 VERSION AC128082.1 GI:21908679
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE Rattus norvegicus.
 ORGANISM Rattus norvegicus

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 178228)

Muzny,D.M., Adams,C., Adio-Oduola,B., Al-Osman,F.R., Allen,C.,
Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barberia,J., Benton,J., Blmage,K., Blankenburg,K., Bonin,D.,
Buck,J., Bowle,S., Bileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devilla,M.L., Davis,C., Day-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homs,I.F., Howard,S., Huber,J., Huily,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Koryah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozad,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Matlindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,S., Ogih,M., Okunou,G.,
Nguyen,N., Nickerson,E., Nwokwkw,S., Ogih,M., Okunou,G.,
Oragunye,N., Oviold,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Qulles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Syatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczek,R., Woodson,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

TITLE
JOURNALREFERENCE
AUTHORSTITLE
JOURNAL

COMMENT

Unpublished
2 (bases 1 to 178228)
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: G2QW
Center clone name: CH230-362C16
Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 116771 bases at least Q40
Consensus quality: 123909 bases at least Q30
Consensus quality: 129462 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 64 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1173	contig of 1173 bp in length
1174	gap of unknown length
1274	contig of 1162 bp in length
2435	gap of unknown length
2535	gap of unknown length
2536	contig of 1032 bp in length
3568	gap of unknown length
3668	contig of 1563 bp in length
5231	gap of unknown length
5331	contig of 1402 bp in length
6732	gap of unknown length
6832	contig of 1158 bp in length
6833	gap of unknown length
7991	contig of 1551 bp in length
8091	gap of unknown length
9642	contig of 1035 bp in length
9742	gap of unknown length
10777	contig of 1029 bp in length
10877	gap of unknown length
11906	contig of 1315 bp in length
12006	gap of unknown length
13321	contig of 1258 bp in length
13421	gap of unknown length
14679	contig of 1073 bp in length
14779	gap of unknown length
15852	contig of 1642 bp in length
15952	gap of unknown length
17594	contig of 1218 bp in length
17694	gap of unknown length
18912	contig of 1134 bp in length
19012	gap of unknown length
20146	contig of 1282 bp in length
20246	gap of unknown length
20246	contig of 1311 bp in length
21557	gap of unknown length
21657	contig of 1059 bp in length
22716	gap of unknown length
22815	contig of 1134 bp in length
23949	gap of unknown length
23950	contig of 1282 bp in length
24050	gap of unknown length
25332	contig of 1186 bp in length
25432	gap of unknown length
26177	contig of 1718 bp in length
26618	gap of unknown length
28436	contig of 2125 bp in length
28536	gap of unknown length
30661	contig of 1503 bp in length
30761	gap of unknown length
32264	contig of 1802 bp in length
32264	gap of unknown length
34165	contig of 1037 bp in length
34265	gap of unknown length
35302	contig of 1880 bp in length
35303	gap of unknown length
35403	contig of 1778 bp in length
37283	gap of unknown length
37383	contig of 1620 bp in length
39161	gap of unknown length
39261	contig of 1456 bp in length
40881	gap of unknown length
40981	contig of 1845 bp in length
42437	gap of unknown length
42537	contig of 1910 bp in length
44382	gap of unknown length
44481	contig of 1918 bp in length
46392	gap of unknown length
46492	contig of 1422 bp in length
48410	gap of unknown length
48510	contig of 1340 bp in length
49932	gap of unknown length
50032	contig of 1340 bp in length
51372	gap of unknown length

```

* 51472 53362: contig of 1891 bp in length
* 53363 53462: gap of unknown length
* 53463 55170: contig of 1708 bp in length
* 55171 55270: gap of unknown length
* 55271 56697: contig of 1427 bp in length
* 56698 56797: gap of unknown length
* 56798 58037: contig of 1240 bp in length
* 58038 58137: gap of unknown length
* 58138 60361: contig of 2224 bp in length
* 60362 60461: gap of unknown length
* 60462 62335: contig of 1874 bp in length
* 62336 62435: gap of unknown length
* 62436 64906: contig of 2471 bp in length
* 64907 65006: gap of unknown length
* 65007 67475: contig of 2469 bp in length
* 67476 67575: gap of unknown length
* 67576 69653: contig of 2078 bp in length
* 69654 73830: gap of unknown length
* 73831 73930: contig of 4077 bp in length
* 73931 77047: gap of unknown length
* 77048 77147: contig of 3117 bp in length
* 77148 80438: gap of unknown length
* 80439 80538: gap of unknown length
* 80539 83532: contig of 2994 bp in length
* 83533 83632: gap of unknown length
* 83633 85754: contig of 2122 bp in length
* 85755 85854: gap of unknown length
* 85855 89203: contig of 3349 bp in length
* 89204 92865: gap of unknown length
* 92866 92965: gap of 3562 bp in length
* 92966 95763: gap of 2798 bp in length
* 95764 95863: gap of unknown length
* 95864 99698: contig of 3833 bp in length
* 99699 99798: gap of unknown length
* 99799 104352: contig of 4554 bp in length
* 104353 104452: gap of unknown length
* 104453 108165: contig of 3713 bp in length
* 108166 108265: gap of unknown length
* 108266 111945: contig of 3680 bp in length
* 111946 112045: gap of unknown length
* 112046 116301: contig of 4256 bp in length
* 116302 116401: gap of unknown length
* 116402 120825: contig of 4424 bp in length
* 120826 120925: gap of unknown length

```

```

Query Match 27.7%; Score 117.6; DB 2; Length 178228;
Best Local Similarity 72.2%; Pred. No. 2.6e-20;
Matches 153; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

```

```

QY 89 CTGGGAGCCCTCCGAGGCCCATCTCACTTCTCCGACTTGGTACAGCAGCCTACCCAT 148
    ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152659 CTCCGAGCCCTCCGAGGCCCATCTCACTTCTCCGACTTGGTACAGCAGCCTACCCAT 152600
QY 149 TTCTGAGATGTTTCCCACTATTCACATGCGTCCCGCAGCCCATTTTGGAGCAGC 208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152599 GTCTGAGAGATTGTCCCACTATTCACATGCGTCCCGCAGCCCATTTTGGAGCAGC 152540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 209 GGTGCGCCAGCAGCAGTCTCTGACCTGACAGGAGCCGAGTTGGAGCCCGGGA 268
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152539 GGTGCGCCAGCAGCAGTCTCTGACCTGACAGGAGCCGAGTTGGAGCCCGGGA 152480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 269 GCAGAGCCAGGAGTGGCCCGGAGTACCAATGGCA 300
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152479 ACATGGCCAGTGGCCCGGAGTACCAATGGCA 152448
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 12
LOCUS AC125909 187998 bp DNA linear HTG 23-JUL-2002
DEFINITION Rattus norvegicus clone CH230-100N7, *** SEQUENCING IN PROGRESS
***, 64 unordered pieces.

```

```

ACCESSION AC125909
VERSION GI:21671507
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 187998)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
          Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayala,M., Banks,T.,
          Barbieri,J., Benton,J., Bimarge,K., Blankenburg,K., Bonnin,D.,
          Bouck,R.J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
          Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
          Cartron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
          Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
          Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
          Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
          Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
          Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
          Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
          Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
          Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
          Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
          Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
          Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
          Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
          Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
          Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
          Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
          Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louieged,H.,
          Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
          Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
          Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
          Miner,G., Miner,Z., Mitchell,T., Mohabadi,K., Morgan,M., Morris,S.,
          Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
          Nguyen,N., Nickerson,E., Nwokoko,S., Ogih,M., Okunou,G.,
          Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
          Peters,L., Pickens,R., Primus,E., Pu,L.L., Qulles,M., Ren,Y.G.,
          Rives,M., Rojas,A., Kojudokan,I., Rolfe,M., Ruiz,S., Saverly,G.,
          Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
          Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
          Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
          Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
          Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,C.,
          Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
          Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,
          Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
          Weinstock,G., and Gibbs,R.
          Direct Submission
          Unpublished
          2 (bases 1 to 187998)
          Worley,K.C.
          Direct Submission
          Submitted (02-JUL-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
          3 (bases 1 to 187998)
          Worley,K.C.
          Direct Submission
          Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
          ----- Genome Center
          Center: Baylor College of Medicine
          Center code: BCM
          Web site: http://www.hgsc.bcm.tmc.edu/
          Contact: hgsc-help@bcm.tmc.edu
          ----- Project Information
          Center project name: GVSa
          Center clone name: CH230-100N7
          ----- Summary Statistics
          Sequencing vector: Plasmid;

```

Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 133860 bases at least Q40
Consensus quality: 142697 bases at least Q30
Consensus quality: 148841 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.bjsgc.bcm.tmc.edu/docs/genbank/draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 64 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1031: contig of 1031 bp in length
1032 1131: gap of unknown length
1132 2334: contig of 1203 bp in length
2335 2434: gap of unknown length
2435 3488: contig of 1054 bp in length
3489 3588: gap of unknown length
3589 4860: contig of 1272 bp in length
4861 4960: gap of unknown length
4961 6021: contig of 1061 bp in length
6022 6121: gap of unknown length
6122 7174: contig of 1053 bp in length
7175 7274: gap of unknown length
7275 8532: contig of 1258 bp in length
8533 8632: gap of unknown length
8633 9611: contig of 1079 bp in length
9611 9712: gap of unknown length
9712 11308: contig of 1497 bp in length
11309 11408: gap of unknown length
11409 12604: contig of 1196 bp in length
12605 12704: gap of unknown length
12705 14114: contig of 1410 bp in length
14115 14214: gap of unknown length
14215 16706: contig of 2492 bp in length
16707 16806: gap of unknown length
16807 17875: contig of 1069 bp in length
17876 17975: gap of unknown length
17976 19654: contig of 1679 bp in length
19655 19754: gap of unknown length
19755 21683: contig of 1929 bp in length
21684 21783: gap of unknown length
21784 22871: contig of 1088 bp in length
22872 22971: gap of unknown length
22972 24666: contig of 1695 bp in length
24667 24766: gap of unknown length
24767 26943: contig of 2177 bp in length
26944 27043: gap of unknown length
27044 29057: contig of 2014 bp in length
29058 29157: gap of unknown length
29158 30973: contig of 1816 bp in length
30974 31073: gap of unknown length
31074 32838: contig of 1765 bp in length
32839 32938: gap of unknown length
32939 34616: contig of 1678 bp in length
34617 34716: gap of unknown length
34717 37198: contig of 2482 bp in length
37199 37298: gap of unknown length
37299 39477: contig of 2179 bp in length
39478 39577: gap of unknown length
39578 41815: contig of 2238 bp in length
41816 41915: gap of unknown length
41916 44488: gap of 2573 bp in length
44489 44588: gap of unknown length
44589 46062: contig of 1474 bp in length
46063 46162: gap of unknown length
46163 48328: contig of 2166 bp in length
48329 48428: gap of unknown length
48429 51395: contig of 2967 bp in length

51396 51495: gap of unknown length
51496 54159: contig of 2664 bp in length
54160 54259: gap of unknown length
54260 56383: contig of 2124 bp in length
56384 56483: gap of unknown length
56484 58779: contig of 2296 bp in length
58780 58879: gap of unknown length
58880 60911: contig of 2032 bp in length
60912 61011: gap of unknown length
61012 63907: contig of 2896 bp in length
63908 64007: gap of unknown length
64009 66202: contig of 2195 bp in length
66203 66302: gap of unknown length
66303 69738: contig of 3436 bp in length
69739 69838: gap of unknown length
69839 72410: contig of 2572 bp in length
72411 72510: gap of unknown length
72511 76219: contig of 3709 bp in length
76220 79029: contig of 2710 bp in length
79030 79129: gap of unknown length
79130 83311: contig of 4182 bp in length
83312 83411: gap of unknown length
83412 86394: contig of 2983 bp in length
86395 86494: gap of unknown length
86495 88582: contig of 2088 bp in length
88583 90972: gap of unknown length
90973 91072: contig of 2290 bp in length
91073 93875: gap of unknown length
93876 93975: gap of unknown length
93976 98771: contig of 4796 bp in length
98772 98872: gap of unknown length
98873 101999: contig of 3128 bp in length
102000 102099: gap of unknown length
102100 105352: contig of 3253 bp in length
105353 105452: gap of unknown length
105453 109873: contig of 4421 bp in length
109874 109973: gap of unknown length
109974 112881: contig of 2908 bp in length
112882 112981: gap of unknown length
112982 117279: contig of 4298 bp in length
117280 117379: gap of unknown length
117380 121634: contig of 4255 bp in length
121635 121734: gap of unknown length
121735 125920: contig of 4186 bp in length
125921 126020: gap of unknown length
126021 129909: contig of 3889 bp in length
129910 130009: gap of unknown length
130010 135723: contig of 5714 bp in length
135724 135823: gap of unknown length

Query Match 19.4% Score 82.4; DB 2; Length 187998;
Best Local Similarity 81.9% Pred. No. 4.3e-11;
Matches 95; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 310 CCGGGGGTCTATGCTCTACCTGACACATCTACATCTACATGACACAGTACTGGGG 369
DB 154396 CTGACGGCTCTGTGACTGACCGGCAATATCTACATATACAAAGGCGCAATGCTGGGG 154455
QY 370 GACACCGGGTCTGAGACCTCCAGCTACCCCGAAGCTCCATATCCCATTTCC 425
DB 154456 GAACACGGGGCCCTGAGACCTCCAGCTCCCGCTGAGCTCCATATCCGACCTCC 154511

RESULT 13
AF429315/c 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS Homo sapiens.
SOURCE

ORGANISM	Homo sapiens
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ATTNORS	1 (bases 1 to 125020)
TITLE	Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S., Igersoll-Ashworth,R.G. & Fleisher,A., Stevanin,G., Brice,A., Potter,N.T., Ross,C.A. and Margolis,R.L. A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2
JOURNAL	Nat. Genet. 29 (4), 377-378 (2001)
MEDLINE	2 (bases 1 to 125020)
PUBMED	21583737
REFERENCE	11694876
ATTNORS	Holmes,S.E., Igersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
TITLE	Direct Submission
JOURNAL	Submitted (05-Oct-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES	Location/Qualifiers
SOURCE	1. 125020
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="16"
	/map="16q24.3; between D16S520 and W-12410"
	/note="Isolated from a patient with Huntington's Disease-like 2 (HD12)"
repeat_region	complement(35581..35746)
	/rpt_type=tandem
gene	/rpt_unit=ctg
	complement(<36507..>36887)
	/gene="JPH3"
mrna	/note="JPH3"
	complement(<36507..>36887)
	/gene="JPH3"
CDS	/product="junctophilin 3"
	complement(<36507..36887)
	/gene="JPH3"
	/note="component of the junctional complex between plasma membrane and endoplasmic reticulum"
	/codon_start=1
	/product="junctophilin 3"
	/protein_id="AAL40941.1"
	/db_xref="GI:17646245"
	/translation="MSSGGRFNDGSGSYCGGWDKANHGVCTGRKGGETGSMSS HGEVLGVTPWPSGNTYQGTWADGKRIGLIESGKKNVYGEWTHGKRGYREKAG NGAKYEETWSNGLDGGTETYSDG"
BASE COUNT	29056 a 32731 c 30696 g 28283 t 4254 others
ORIGIN	
Query Match	9.48; Score 40; DB 9; Length 125020;
Best Local Similarity	11.88; Pred. No. 5.6;
Matches	36; Conservative 166; Mismatches 97; Indels 5; Gaps 2;
OY	11 GCCACTGCCGCCAGATGATCGAGTGCCTCTCTCAAGAGGGGTCGCCAGAGGAGGACC 70
Db	51038 SCCTKSKTSCGMGRSSCTCTGCGMCCGMMSSCCASRGCK-STKMGSMRCRRRMS 50980
OY	71 CAATCTGTAGCTGGAAGCTGGAGCCCTCCGAAGGCCATCATCACTTCCCTGACTTGGT 130
Db	50979 CMWNSMKSYSMAWYKSSARBRGMKCYTSCMKYIYSHBKSNDHSYKYSRYBYWT 50920
OY	131 ACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCACTATCACCCTGCGGCTCCCGCAGC 190
Db	50919 CMKMSKMSWSDSHMRDMYDKMKRKYSSGTRGHBHSAWMCKMSRSDSGSG--- 50864
OY	191 CCCAGTTTGGAGGAGGGGTGCCGACACACAGAGCTCTCTGAGACTGCACAGGAGACC 250
Db	50863 MYHMKSYKSSKSMYISGMKSSYBMDCKSMYTSBSKSVMBBRMSGSYBRYKCAVM 50804
OY	251 GCAGTTGGAACCCGGGAGAGAGCAGCAGTGGCCACGATGCATTCATTCATGATGATG 310
Db	50803 RRYBHRSTRSMYMGGRKSCWYDVSYSRRSRVHTTSMSSMRRCRGAGAMANSACAKMCSV 50744
OY	311 CGGC 314

```

Db  50743 MTGM 50740

RESULT 14
LOCUS      AF429315
DEFINITION Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION  AF429315
VERSION     AF429315.1
KEYWORDS   GI:17646244
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 125020)
AUTHORS     Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
            Ingersoll-Ashtworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
            Potter,N.T., Ross,C.A. and Margolis,R.L.
TITLE       A repeat expansion in the gene encoding junctophilin-3 is
            associated with Huntington disease-like 2
JOURNAL     Nat. Genet. 29 (4), 377-378 (2001)
MEDLINE     11594876
PUBMED      11583737
REFERENCE   2 (bases 1 to 125020)
AUTHORS     Holmes,S.E., Ingersoll-Ashtworth,R.G., Ross,C.A. and Margolis,R.L.
TITLE       Direct Submission
JOURNAL     Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
            Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
SOURCE      Location/Qualifiers
            1..125020
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="16"
            /map="16q24.3; between D16S520 and WI-12410"
            /note="Isolated from a patient with Huntington's
            Disease-Like 2 (HDL2)"
            repeat_region complement(3581..35746)
            /rpt_type=tandem
            /rpt_unit=ctg
            gene complement(<36507..36887)
            /note="JPH3"
            mRNA complement(<36507..36887)
            /note="JPH3"
            /gene="JPH3"
            /product="junctophilin 3"
            complement(<36507..36887)
            /gene="JPH3"
            /note="Component of the junctional complex between plasma
            membrane and endoplasmic reticulum"
            /codon_start=1
            /product="junctophilin 3"
            /protein_id="A1A10941.1"
            /db_xref="GI:17646245"
            /translation="MSSGGRRNFDGSGYSCGQWEDGAHGHVCTGPKGQGETGSMSS
            HGEFLVLTWPSGNTYQGTWAGKRGIGLESKGVKVVYKGEWTHGFRGVRECAAG
            NGAYEESTWSKGLDDGXTLEYSDG"
BASE COUNT  29056 a 32731 c 30696 g 28283 t 4254 others
ORIGIN
Query Match 9.0%; Score 38.4; DB 9; Length 125020;
Best Local Similarity 9.5%; Pred. No. 15;
Matches 39; Conservative 196; Mismatches 172; Indels 5; Gaps 1;

QY 8 AGAGCCACTGCCCCAGAGATCTCAGANTCGCTGCTCAAGAGCGCTCCGACAGGAGAGG 67
    111::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17290 GGMSSSYMKMKKSSWRMSCYWKCCWMKMKCCMMRSMSSRMSYYMTASMKSSSRGCTC 17349
    111::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 68 ACCCAATCTGTAGCTGGAAGCTGGAGACCTTCGAGAGCCCATCATCTTCCCTGACTT 127
    111::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17350 TRCYWCMKSKGCKYKYYMMRRKRKKKKMKWSMCMSMMGMSASRSSCKYKMRCSMSS 17409
    111::: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 128 GGTACAGCCACTGTACCACTTTCTGAGAGATGTTTTCCCAAGTATCCACTGGGCTCCCCGC 187

```


GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:05:15 ; Search time 73.9741 Seconds

(without alignments)
12938.300 Million cell updates/sec

Title: US-09-917-372-2_COPY_853_1277

Perfect score: 425

Sequence: 1 atccatgagccactgcc.....aacctcatccaccatccc 425

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	393	92.5	1594	24	ABQ55003
2	393	92.5	2136	24	ABN95637
3	393	92.5	2136	24	ABK64502
4	393	92.5	2136	24	ABL64078
5	393	92.5	2136	24	ABL64384
6	253.8	59.7	281	24	ABL82623
7	251.2	59.1	289	24	ABL82688
8	36.2	8.5	450	23	ABV38197
9	35.2	8.3	333	20	AAV72224

10	34.2	8.0	268	13	AAQ25135	DRI. Synthetic.
11	33.6	7.9	33675	24	AAD31202	Oryza sativa gcpe
12	33.2	7.8	955	22	AAS01024	Sugarcane plant ge
13	33.2	7.8	1007	22	AAS01022	Sugarcane plant ge
14	33	7.8	1057	21	AAZ61623	CDNA encoding rat
15	33	7.8	1057	22	AAQ99556	Skin cell CDNA, SE
16	33	7.8	1057	24	ABL34708	Scapulariopsis bre
17	33	7.8	1722	18	AA784865	S. brevicaulis bet
18	33	7.8	1839	20	AAZ28415	Skin cell CDNA, SE
19	33	7.8	1890	22	AAQ99799	Rat CDNA isolated
20	33	7.8	1880	24	ABL34951	S. brevicaulis bet
21	33	7.8	1894	20	AAZ28424	CDNA encoding rat
22	33	7.8	2004	21	AAZ61757	Skin cell CDNA, SE
23	33	7.8	2004	22	AAQ99690	Rat CDNA isolated
24	33	7.8	2004	24	ABL34842	Drosophila melanog
25	33	7.8	13085	23	ABL03680	Human nervous syst
26	33	7.8	16818	22	ABA15470	Human nervous syst
27	33	7.8	16818	22	ABA16242	Human nervous syst
28	33	7.8	16818	22	ABA18914	Human gene signatu
29	32.8	7.7	186	16	AAZ26531	Human secreted pro
30	32.8	7.7	376	19	AAV59700	DNA encoding novel
31	32.6	7.7	1527	23	AA668546	DNA encoding novel
32	32.6	7.7	1620	23	AA668547	DNA encoding novel
33	32.6	7.7	2178	23	AA586106	Drosophila melanog
34	32.6	7.7	3112	23	ABL05387	Drosophila melanog
35	32.6	7.7	5721	23	ABL05386	Human MDRP encodin
36	32.4	7.6	649	24	ABQ72658	Human MDRP encodin
37	32.4	7.6	1056	21	AAZ51683	Zea mays DNA fragm
38	32.4	7.6	1276	24	ABQ72500	Human MDRP encodin
39	32.4	7.6	1677	23	AA588638	DNA encoding novel
40	32.4	7.6	4677	23	AA589696	DNA encoding novel
41	32.4	7.6	4677	23	AA592991	DNA encoding novel
42	32.2	7.6	785	24	ABK39773	CDNA encoding lung
43	32.2	7.6	2861	12	AAQ14383	Clone 1 encoding I
44	32.2	7.6	2861	21	AAZ21133	Human low adenosin
45	32.2	7.6	2861	21	AAZ35011	Human adenosine re

ALIGNMENTS

RESULT 1	ABQ55003	ABQ55003 standard; cDNA; 1594 BP.
ID	ABQ55003	
AC	ABQ55003;	
XX		
DT	22-AUG-2002 (first entry)	
XX		
DE	Human ovarian antigen HSABJ44 cDNA, SEQ ID NO:883.	
XX		
KW	Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;	
KW	ovarian cancer; breast cancer; tumour; reproductive system disorder;	
KW	infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;	
KW	PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;	
KW	inflammatory condition; immune disorder; blood disorder;	
KW	cardiovascular disorder; respiratory disorder; neurological disorder;	
KW	gastrointestinal disorder; urinary system disorder; drug screening;	
KW	gene therapy; chromosome mapping; forensic analysis;	
KW	antibody preparation; cytostatic; immunomodulatory; neuroprotective;	
KW	antiinflammatory; gynaecological; reproductive; chromosome 12p13;	
KW	gene; ss.	
OS	Homo sapiens.	
XX		
PN	WO200200677-A1.	
XX		
PD	03-JAN-2002.	
XX		
PF	07-JUN-2001: 2001MO-US18569.	
XX		
PR	07-JUN-2000: 2000US-209467P.	
XX		

XX	(HUMA-)	HUMAN GENOME SCI INC.	
PI	Blrse CE, Rosen CA;		
DR	WPI, 2002-147878/19.		
DR	P-PSDB; ABP41926.		
XX			
PT	Isolated nucleic acid molecules encoding novel ovarian polypeptides,		
PT	useful in the prevention, treatment and diagnosis of cancer (e.g.,		
PT	ovarian cancer), immune disorders, cardiovascular disorders and		
PT	neurological diseases -		
XX			
PS	Claim 1; SEQ ID No 883; 2922pp; English.		
XX			
CC	The invention relates to 2175 novel human ovarian antigens (ABP41054-		
CC	ABP413228) and to cDNAs encoding them (AB054131-AB056305), and also		
CC	encompasses polypeptides 90% identical and polynucleotides 95% identical		
CC	to the sequences of the invention. The invention additionally relates to		
CC	recombinant vectors and host cells comprising human ovarian antigen		
CC	polynucleotides, antibodies against human ovarian antigens, and the use		
CC	of ovarian antigen polynucleotides and polypeptides in diagnosing,		
CC	treating, prognosing or preventing various ovary and/or breast-related		
CC	disorders. Such conditions include ovarian cancer and breast cancer, and		
CC	metastatic tumours of ovarian or breast origin, reproductive system		
CC	disorders (e.g., infertility, disorders of pregnancy, anovulation,		
CC	polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine		
CC	disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic		
CC	shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and		
CC	varicella), immune disorders (e.g., congenital and acquired		
CC	immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus)		
CC	blood-related disorders (e.g., anaemia), cardiovascular disorders,		
CC	respiratory disorders, neurological disorders, gastrointestinal disorders		
CC	and urinary system disorders. Ovarian antigen polypeptides and		
CC	polynucleotides may also be used in screening for compounds which		
CC	modulate ovarian antigen expression or activity. The polynucleotides may		
CC	further be used for gene therapy, chromosome mapping, in the		
CC	identification of individuals and in forensic analysis, and the		
CC	polypeptides may be used as food additives or to prepare antibodies		
CC	useful in disease diagnosis, drug targeting and phenotyping. The present		
CC	sequence represents cDNA encoding a human ovarian antigen of the		
CC	invention.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp://wipo.int/pub/published_pot_sequences .		
CC			
XX			
SO	Sequence 1594 BP; 391 A; 488 C; 434 G; 280 T; 1 other;		
	Query Match	92.5%; Score 393; DB 24; Length 1594;	
	Best Local Similarity 100.0%; Pred. No. 2.8e-100;		
	Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps		
QY	33	GGATGGCTGCTCAGAGGCGTCCCGAGGAGAGGAGACCAATCTCTGACTGGAACTGG	92
DB	395	GGATGGCTGCTCAGAGGCGTCCCGAGGAGAGGAGACCAATCTCTGAGTGGAACTGG	454
QY	93	GAGCTCCGAGAGGCCATCATCTTCCTCGACTTGGTAAAGCACAGTCATCCATTTC	152
DB	455	GAGCTCCGAGAGGCCATCATCTTCCTCGACTTGGTAAAGCACAGTCATCCATTTC	514
QY	153	GGAGATGTTTCCCAAGTATCCACTGGGCTCCCGACGCCCACTTTTGAAGAGGAGG	212
DB	515	GGAGATGTTTCCCAAGTATCCACTGGGCTCCCGACGCCCACTTTTGAAGAGGAGG	574
QY	213	CCGCAACAGCAGAGTCTCTGAGCTATCCAGGAGAGGCCGAGTGGAAACCCGGGAGCAG	272
DB	575	CCGCAACAGCAGAGTCTCTGAGCTATCCAGGAGAGGCCGAGTGGAAACCCGGGAGCAG	634
QY	273	AGCCAGTGGGCCAGAGTACCAATGGCAATTCATCCAGCGCGGCTATAGACTATCACT	332
DB	633	AGCCAGTGGGCCAGAGTACCAATGGCAATTCATCCAGCGCGGCTATAGACTATCACT	694
QY	333	GGCAACATCTACATCTACAAATGAGACACTAGTGGGGGAGACACCGGGTCTCGAGACCTC	392

Db	695	GGCAACATCTACATCTACATCAATGACACACTACTGGGGGGGACACACGGGCTCTGGAGACCTC	754
QY	393	CCAGCTACCCCCGAACTCTCATACCCATTCCC	425
Db	755	CCAGCTACCCCCGAACTCTCATACCCATTCCC	787
RESULT 2			
ID	ABN95637	standard; DNA; 2136 BP.	
AC	ABN95637;		
XX			
XX	13-AUG-2002 (first entry)		
DE	Gene #2135 used to diagnose liver cancer.		
XX			
KW	Gene: liver cancer; ds; hepatocellular carcinoma; hepatotropic;		
XX	metastatic liver tumour; cytostatic; expression profile; disease state;		
KW	disease progression; drug toxicity; drug efficacy; drug metabolism.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200229103-A2.		
PD	11-APR-2002.		
XX			
PF	02-OCT-2001; 2001WO-US30589.		
XX			
PR	02-OCT-2000; 2000US-237054P.		
XX			
PA	(GENE-) GENE LOGIC INC.		
PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;		
DR	WPI; 2002-426119/45.		
XX			
PT	Diagnosing and detecting the progression of liver cancer,		
XX	hepatocellular carcinoma or metastatic liver tumor in a patient,		
PT	involves detecting the level of expression of two or more genes in a		
XX	liver tissue sample		
PS	Claim 1; SEQ ID NO 2135; 298pp; English.		
XX			
CC	The invention relates to a novel method for diagnosing and detecting the		
CC	progression of liver cancer, hepatocellular carcinoma or metastatic liver		
CC	tumour in a patient, and differentiating metastatic liver cancer from		
CC	hepatocellular carcinoma in a patient, involving detecting the level of		
CC	expression of two or more genes represented in ABN93503-ABN97455 in a		
CC	tissue sample. The method of the invention has hepatotropic, and		
CC	cytostatic activity. The method is useful for diagnosing and detecting		
CC	the progression of liver cancer, hepatocellular carcinoma and metastatic		
CC	liver carcinoma in a patient. The method is useful for identifying		
CC	expression profiles which serve as useful diagnostic markers as well as		
CC	markers that can be used to monitor disease states, disease progression,		
CC	drug toxicity, drug efficacy and drug metabolism.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX			
SEQ	Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other:		
Query Match	92.5%; Score 393; DB 24; Length 2136;		
Best Local Similarity	100.0%; Pred. No. 3,1e-100;		
Matches 393; Conservative	0; Mismatches 0; Indels 0; Gaps 0		
QY	33	GGATGCTCTCTCAAGAGGCGTCGCGAGGAGAGGAGCCCAATCCTTACTGGAACCTGG	92
Db	943	GGATGCTCTCTCAAGAGGCGTCGCGAGGAGGAGCCCAATCCTTACTGGAACCTGG	1002
QY	93	GAGCCTCCGAAAGCCATCATCTCTCCCTGACTTGGTACAGCACTGTACCATTTCT	152
Db	1003	GAGCCTCCGAAAGCCATCATCTCTCCCTGACTTGGTACAGCACTGTACCATTTCT	1062

OY 153 GGAGATGTTTCCCACTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGCGAGGGGTG 212
 |||||||
 DB 1063 GGAGATGTTTCCCACTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGCGAGGGGTG 1122
 OY 213 CCGCAACAGCAGAGTCTCTGACCTGACCGAGGAGCCGAGTTGGACCCGGGGAGCAG 272
 |||||||
 DB 1123 CCGCAACAGCAGAGTCTCTGACCTGACCGAGGAGCCGAGTTGGACCCGGGGAGCAG 1182
 OY 273 AGCCAGTGGCCGACGGTACCAATGGCATTCATGTCACGGCGGGTCTATGACTATCACT 332
 |||||||
 DB 1183 AGCCAGTGGCCGACGGTACCAATGGCATTCATGTCACGGCGGGTCTATGACTATCACT 1242
 OY 333 GGCACATGTCATCTACATGAGACCACTAGTGGGGGAGCACCCGGGCTCTGGAGACCTC 392
 |||||||
 DB 1243 GGCACATGTCATCTACATGAGACCACTAGTGGGGGAGCACCCGGGCTCTGGAGACCTC 1302
 OY 393 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 425
 |||||||
 DB 1303 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 1335

RESULT 3

ABL64502
 ID ABL64502 standard; DNA; 2136 BP.

AC ABL64502;

DT 18-JUN-2002 (first entry)

XX Human benign prostatic hyperplasia gene #397.

XX Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.

OS Homo sapiens.

PN WO200212440-A2.

PD 14-FEB-2002.

PF 07-AUG-2001; 2001WO-US24708.

PR 07-AUG-2000; 2000US-223323P.

PR 05-JUN-2001; 2001US-0873319.

XX (GENE-) GENE LOGIC INC.

PA (MISB) JAPAN TOBRACO INC.

PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;

DR WPI: 2002-257476/30.

XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by

PT detecting expression levels of one or more genes in prostate cells from

PT patient that are differentially regulated compared to normal prostate

PS cells

XX Disclosure; Page 229-230: 444pp: English.

XX The invention relates to a method of diagnosing (I) the onset or

CC progression of benign prostatic hyperplasia (BPH), or screening (II) for

CC or identifying an agent that modulates the onset or progression of BPH.

CC The method is based on changes in gene expression in BPH tissue isolated

CC from patients exhibiting different clinical states of prostate

CC hyperplasia as compared to normal prostate tissue. (I) comprises

CC detecting the expression levels of one or more genes in prostate cells

CC from the subject that are differentially regulated compared to normal

CC prostate cells; (II) comprises preparing a first gene expression profile

CC of BPH cells or BPH-like cell population, exposing the cells to the

CC agent, preparing a second gene expression profile of the agent exposed

CC cells, and comparing the first and second gene expression profiles.

CC (I) is useful for diagnosing the onset or progression of BPH. (II) is

CC useful for identifying an agent that modulates the onset or progression

CC of BPH. The methods are useful to present information identifying
 CC the expression level in a tissue or cells, by comparing the expression
 CC level of genes given in the specification, in the tissue or cells to the
 CC level of expression of gene in the database, and displaying the
 CC expression levels of at least one gene in the tissue or cell sample
 CC compared to the expression level in BPH. Agents using (II) are useful for
 CC treating BPH or prostate cancer. ABL64106-ABL64860 represent human
 CC benign prostatic hyperplasia gene sequences of the invention.

XX Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other:

Query Match 92.5%; Score 393; DB 24; Length 2136;

Best Local Similarity 100.0%; Pred. No. 3.1e-100;

Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 GGATCGTCTCTCAAGAGCGCTCCGACGAGGAGAGGAGCCCAATCTGTAGCTGGAAGCTGG 92

DB 943 GGATCGTCTCTCAAGAGCGCTCCGACGAGGAGAGGAGCCCAATCTGTAGCTGGAAGCTGG 1002

OY 93 GAGCCCTCGAAGGCGCATCATCTTCCCTGACTTGTGACGACCTGCTACCATTTCT 152

DB 1003 GAGCCCTCGAAGGCGCATCATCTTCCCTGACTTGTGACGACCTGCTACCATTTCT 1062

OY 153 GGAGATGTTTCCCACTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGCGAGGGGTG 212

DB 1063 GGAGATGTTTCCCACTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGCGAGGGGTG 1122

OY 213 CCGCAACAGCAGAGTCTCTGACCTGACCGAGGAGCCGAGTTGGACCCGGGGAGCAG 272

DB 1123 CCGCAACAGCAGAGTCTCTGACCTGACCGAGGAGCCGAGTTGGACCCGGGGAGCAG 1182

OY 273 AGCCAGTGGCCGACGGTACCAATGGCATTCATGTCACGGCGGGTCTATGACTATCACT 332

DB 1183 AGCCAGTGGCCGACGGTACCAATGGCATTCATGTCACGGCGGGTCTATGACTATCACT 1242

OY 333 GGCACATGTCATCTACATGAGACCACTAGTGGGGGAGCACCCGGGCTCTGGAGACCTC 392

DB 1243 GGCACATGTCATCTACATGAGACCACTAGTGGGGGAGCACCCGGGCTCTGGAGACCTC 1302

OY 393 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 425

DB 1303 CCAGCTACCCCGCAACCTCCATACCCCATTTCC 1335

RESULT 4
 ABL64078
 ID ABL64078 standard; DNA; 2136 BP.

XX ABL64078;

DT 15-MAY-2002 (first entry)

XX Breast cancer related gene sequence SEQ ID NO:2415.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

XX cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;

XX gene; ds.

XX Homo sapiens.

PN WO200194629-A2.

PD 13-DEC-2001.

PF 30-MAY-2001; 2001WO-US10838.

PR 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 PA (AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI: 2002-188264/24.
 PT
 PT Screening for anti-neoplastic agent involves exposing cells to a
 chemical agent to be tested for anti-neoplastic activity, and
 determining a change in expression of a gene of a signature gene set
 XX
 PS Claim 1; SEQ ID 2415; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 anti-neoplastic agent. The method involves exposing cells to a chemical
 agent to be tested for anti-neoplastic activity, determining a change in
 expression of at least one gene (I) of a signature gene set, where (I)
 comprises a sequence (S) selected from 8447 sequences (given in ABL6164
 to ABL70110), or is at least 95% identical to (S), where a change in
 expression is indicative of anti-neoplastic activity. (I) has cytosolic
 activity and can be used in gene therapy. M1 can be used for screening
 an anti-neoplastic agent, and can be used for producing a product which
 is the data collected with respect to the anti-neoplastic agent as a
 result of M1, and the data is sufficient to convey the chemical
 structure and/or properties of the agent. M1 can be used in the
 treatment of cancer such as colon, breast, stomach, lung, thyroid,
 oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX
 SO Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

QY 33 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGACCAATCTGTACTGGAACCTGG 92
 Db 943 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGACCAATCTGTACTGGAACCTGG 1002
 QY 93 GAGCCTCCGAAAGGCCCATTCATCTCCCTGACTGTGGTACAGCCACTGCTACCATTTCT 152
 Db 1003 GAGCCTCCGAAAGGCCCATTCATCTCCCTGACTGTGGTACAGCCACTGCTACCATTTCT 1062
 QY 153 GGAGATTTTCCCGCAGTATCCACTGGGCTCCCGGAGCCCACTTTTGGAGGAGGGGTG 212
 Db 1063 GGAGATTTTCCCGCAGTATCCACTGGGCTCCCGGAGCCCACTTTTGGAGGAGGGGTG 1122
 QY 213 CCGCAACAGCAGAGTCTCTGACCTTGACCCAGGAGCCGAGTTGAAACCCGGGAGCAG 272
 Db 1123 CCGCAACAGCAGAGTCTCTGACCTTGACCCAGGAGCCGAGTTGAAACCCGGGAGCAG 1182
 QY 273 AGCCAGGTGGCCCGCAGGTACCATGSCATTCATGTCAACCGGGGCTGTATGACTATCACT 332
 Db 1183 AGCCAGGTGGCCCGCAGGTACCATGSCATTCATGTCAACCGGGGCTGTATGACTATCACT 1242
 QY 333 GGCAACATCTACATCTACATGACGACGATGCGGGGGGACACCGGGTCTGAGAGACTC 392
 Db 1243 GGCAACATCTACATCTACATGACGACGATGCGGGGGGACACCGGGTCTGAGAGACTC 1302
 QY 393 CCAGCTACCCCGCAACCTCCATACCCCATTCGCC 425
 Db 1303 CCAGCTACCCCGCAACCTCCATACCCCATTCGCC 1335
 RESULT 5
 ABL64384
 ID ABL64384 standard; DNA; 2136 BP.
 AC
 XX ABL64384;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Stomach cancer related gene sequence SEQ ID NO:2721.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; anti-neoplastic; Wilms' tumour; adenocarcinoma;
 gene; ds.
 XX
 OS Homo sapiens.
 XX
 PM WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-2095473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.

PR	27-SEP-2000;	2000US-235863P.	
PR	28-SEP-2000;	2000US-236028P.	
PR	28-SEP-2000;	2000US-236032P.	
PR	28-SEP-2000;	2000US-236033P.	
PR	28-SEP-2000;	2000US-236034P.	
PR	28-SEP-2000;	2000US-236109P.	
PR	28-SEP-2000;	2000US-236111P.	
PR	29-SEP-2000;	2000US-236842P.	
PR	29-SEP-2000;	2000US-236891P.	
PR	02-OCT-2000;	2000US-237172P.	
PR	02-OCT-2000;	2000US-237173P.	
PR	02-OCT-2000;	2000US-237278P.	
PR	02-OCT-2000;	2000US-237294P.	
PR	02-OCT-2000;	2000US-237295P.	
PR	02-OCT-2000;	2000US-237316P.	
PR	03-OCT-2000;	2000US-237425P.	
PR	03-OCT-2000;	2000US-237598P.	
PR	03-OCT-2000;	2000US-237604P.	
PR	03-OCT-2000;	2000US-237606P.	
PR	03-OCT-2000;	2000US-237608P.	
PR	01-NOV-2000;	2000US-244867P.	
PR	01-NOV-2000;	2000US-245084P.	
PA	(AVAL-)	AVALON PHARM.	
XX			
PI	Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;		
PI	Soppet DR, Weaver Z;		
XX			
DR	WPI; 2002-188264/24.		
XX			
PT	Screening for anti-neoplastic agent involves exposing cells to a		
PT	chemical agent to be tested for anti-neoplastic activity, and		
PT	determining a change in expression of a gene of a signature gene set		
XX			
PS	Claim 1; SEQ ID 2721; 44pp; English.		
CC	The present invention describes a method (M1) for screening for an		
CC	anti-neoplastic agent. The method involves exposing cells to a chemical		
CC	agent to be tested for anti-neoplastic activity, determining a change in		
CC	expression of at least one gene (I) of a signature gene set, where (I)		
CC	comprises a sequence (S) selected from 8447 sequences (given in AB01664		
CC	to AB170110), or is at least 95% identical to (S), where a change in		
CC	expression is indicative of anti-neoplastic activity. (I) has cytostatic		
CC	activity and can be used in gene therapy. M1 can be used for screening		
CC	an anti-neoplastic agent, and can be used for producing a product which		
CC	is the data collected with respect to the anti-neoplastic agent as a		
CC	result of M1, and the data is sufficient to convey the chemical		
CC	structure and/or properties of the agent. M1 can be used in the		
CC	treatment of cancer such as colon, breast, stomach, lung, thyroid,		
CC	adenocarcinoma, carcinoma, kidney, prostate or pancreatic cancer,		
CC	adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,		
CC	infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine		
CC	carcinoma, papillary carcinoma and Wilms' tumor.		
XX			
XX	Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other:		
XX			
Query Match	92.5%; Score 393; DB 24; Length 2136;		
Best Local Similarity	100.0%; Pred. No. 3.1e-100; Length 2136;		
Matches	393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	33 GGATGCTGCTCAGAGGCGCTCCGAGGAGAGGAGACCCATCTGTAAGCTGG 92		
DB	943 GGATCGCTGCTCAGAGGCGCTCCGAGGAGAGGAGACCCATCTGTAAGCTGG 1002		
OY	93 GAGCCTCCGAGGCGCCATCTGCTGCTGCTGTAAGGACCTGTAAGCTGGT 152		
DB	1003 GAGCCTCCGAGGCGCCATCTGCTGCTGCTGTAAGGACCTGTAAGCTGGT 1062		
OY	153 GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGAGGCGCCAGTTTGGAGCAGGGGTG 212		
DB	1063 GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGAGGCGCCAGTTTGGAGCAGGGGTG 1122		
OY	213 CCGCACAAGCAGATGCTCTGTGACCTGACCAAGGAGCCGCACTTGGAAACCGGGAGCAG 272		

Db	1123	CGGACACGCGAGAGTCTCTTGAGACTTACACGAGGAGCGGAGTTGGAAACCCGGGAGACAG	1182
Oy	273	AGCCAGGTGGGCGCCACGGTACCAATGAGCATTCATGTACCGCGGCGGTCTATGATATATACAT	332
Db	1183	AACCGAGTGGGCCCAACGGAGACCAATGGCATTCATGTACCGCGGCGGTCTATGATATATACAT	1242
Oy	333	GGCAACATCTCATCTACTACATGAGCAGCAGTACTGGGGGGGACCCACCGGGTCTCTGAGACCTC	392
Db	1243	GGCAACATCTCATCTACTACATGAGCAGCAGTACTGGGGGGGACCCACCGGGTCTCTGAGACCTC	1302
Oy	393	CCAGCTACCCCGGAACCTCCATACCCTATCC	425
Db	1303	CCAGCTACCCCGGAACCTCCATACCCTATCC	1335
RESULT 6			
ABL82623			
XX	ID	ABL82623 standard; cDNA; 281 BP.	
XX	AC	ABL82623;	
XX	DT	17-MAY-2002 (first entry)	
XX	DE	Human ovarian cancer related cDNA clone SEQ ID NO:5601.	
XX	KW	Human; ovarian cancer; ovarian tumour; cystostatic; gene; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO200192581-A2.	
XX	PD	06-DEC-2001.	
XX	PF	29-MAY-2001; 2001WO-US17756.	
XX	PR	26-MAY-2000; 2000US-207484P.	
XX	PA	(CORI-) CORIXA CORP.	
XX	PI	Algate PA, Harlocker SL, Jones R;	
XX	DR	WPI: 2002-122075/16.	
XX	PT	Composition for therapy and diagnosis of ovarian cancer comprising	
XX	PT	polypeptide of a ovarian tumor polypeptide, polynucleotide encoding	
XX	PT	polypeptide, antibody specific to polypeptide or T cell expressing	
XX	PT	polypeptide -	
XX	PS	Claim 1; SEQ ID 5601; 489bp; English.	
XX	CC	The present invention describes a composition (I) comprising: carriers	
XX	CC	and immunostimulants; and a polypeptide (II) of a ovarian tumour	
XX	CC	polypeptide encoded by a polynucleotide (III) having a cDNA sequence	
XX	CC	(S1) from the 10912 nucleotide sequences as given in ABL77023 to	
XX	CC	AB187934, (III) encoding (II) having a sequence (S2), a T cell	
XX	CC	population of (II), or antigen presenting cells that express (II).	
XX	CC	(I) has cytostatic activity. An oligonucleotide (IV) that hybridises to	
XX	CC	(S1) can be used for detecting ovarian cancer in a patient's biological	
XX	CC	sample preferably serum or ovarian tissue. The method comprises	
XX	CC	contacting a biological sample from a patient with (IV), detecting the	
XX	CC	amount of polynucleotide hybridising to (IV) and comparing the amount to	
XX	CC	a predetermined cutoff value and thereby detecting ovarian cancer in the	
XX	CC	patient, where the amount of polynucleotide hybridising to (IV) is	
XX	CC	detected preferably by polymerase chain reaction (PCR). (I) comprising	
XX	CC	(III) and/or (II) is useful for stimulating and/or expanding T cells	
XX	CC	specific for an ovarian tumour protein comprising contacting T cells	
XX	CC	with (III) or (II). (III) is useful in design and preparation of	
XX	CC	ribozyme molecules for inhibiting expression of the tumour polypeptides	
XX	CC	and proteins in tumour cells; and to isolate a full length gene from a	
XX	CC	suitable library e.g., a tumour cDNA library using well known	
XX	CC	techniques.	

Sequence 281 BP; 58 A; 94 C; 77 G; 52 T; 0 other;

Query Match 59.7%; Score 253.8; DB 24; Length 281;
 Best Local Similarity 97.1%; Pred. No. 1.7e-61;
 Matches 269; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 144 CCATTTCCTGGAGATGTTTCCCGATCCACACGAGGCTCCCGCCAGGCC-CAGTTTGA 202
 11
 DB 5 CCACTTCTGGAGATGTTTCCCGATCCACACGAGGCTCCCGCCAGGCCAGTTTGA 64

QY 203 GGCAGGGGTGCGCAGACAGAGATCTCTGACCTGACGAGGAGCGCGATTGNAAC 262
 65 GGCAGGGGTGCGCAGACAGAGATCTCTGACCTGACGAGGAGCGCGATTGNAAC 124

QY 263 CGGGAGACAGAGCCAGGTGGCCCGACGCTACCATGCTTCAATGTCACCGGGGTAT 322
 125 CGGGAGACAGACTGAGTGGCCCGACGCTACCATGCTTCAATGTCACCGGGGTAT 184

QY 323 GACATATCTGGACATCTACATCTACATGACAGTACTGGGGGAGCACCGGGTCC 382
 185 GACATATCTGGACATCTACATCTACATGACAGTACTGGGGGAGCACCGGGTCA 244

QY 383 TGGAGACCTCCACGCTACCCCGGACCTCCATACCC 419
 245 TGGAGACCTCCACGCTACCCCGGACCTCCATACCC 281

RESULT 7
 ABL82688/c
 ID ABL82688 standard; cDNA; 289 BP.
 XX ABL82688;
 AC
 DB 17-MAY-2002 (first entry)
 DE Human ovarian cancer related cDNA clone SEQ ID NO:5666.
 XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
 KW
 XX Homo sapiens.
 OS
 PN MO200192581-A2.
 PD 06-DEC-2001.
 XX 29-MAY-2001; 2001WO-US17756.
 PF 26-MAY-2000; 2000US-207484P.
 PR (CORI-) CORIXA CORP.
 XX (CORI-) CORIXA CORP.
 PA
 XX Algate PA, Harlocker SL, Jones R;
 PI
 DR WPI; 2002-122075/16.
 XX
 PT Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide
 PS Claim 1; SEQ ID 5666; 489pp; English.
 PS
 XX The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polynucleotide (II) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC Population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 CC (SI) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to

a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.
 CC
 XX Sequence 289 BP; 49 A; 75 C; 93 G; 72 T; 0 other;

Query Match 59.1%; Score 251.2; DB 24; Length 289;
 Best Local Similarity 98.5%; Pred. No. 9.1e-61;
 Matches 264; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 155 AGATGTTTCCCGATGATCCACGAGTGGCTCCCGCCAGCCCGGATTTGAGAGCGAGGGTGC 214
 289 AGATGTTTCCCGATGATCCACGAGTGGCTCCCGCCAGCCCGGATTTGAGAGCGAGGGTGC 230

QY 215 GCACAGCAGAGA-GTCTCTGACCTGACGAGGAGCGGAGTTGNAACCGGGAGCAGA 273
 DB 229 GCACAGCAGAGATGCTCTGACCTGACGAGGAGCGGAGTTGNAACCGGGAGCAGA 170

QY 274 GCCAGGTGGCCCGACGCTACCATGCTTCAATGTCACCGGGGTATGACTATCATCG 333
 DB 169 GCCAGGTGGCCCGACGCTACCATGCTTCAATGTCACCGGGGTATGACTATCATCG 110

QY 334 GCACATCTACATCTACATGACAGTACTGGGGGAGCACCGGGTCTGAGACCTCC 393
 DB 109 GCACATCTACATCTACATGACAGTACTGGGGGAGCACCGGGTCTGAGACCTCC 50

QY 394 CAGCTACCCCGGACCTCCATACCCCAT 421
 DB 49 CAGCTACCCCGGACCTCCATACCCCAT 22

RESULT 8
 ABV38197
 ID ABV38197 standard; cDNA; 450 BP.
 XX ABV38197;
 AC
 DB 16-SEP-2002 (first entry)
 DE Human prostate expression marker cDNA 38188.
 XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX pharmacogenomic marker; gene; ss.
 KW
 XX Homo sapiens.
 OS
 PN WO200160860-A2.
 PD 23-AUG-2001.
 XX 20-FEB-2001; 2001WO-US05171.
 PF 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 DR WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
PS Claim 1; Page 7795; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 450 BP; 90 A; 130 C; 127 G; 103 T; 0 other;
Query Match 8.5%; Score 36.2; DB 23; Length 450;
Best Local Similarity 51.6%; Pred. No. 1.9;
Matches 83; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 30 TCAGATCGCTGCTCAGAGGCGTCCGAGAGGAGAGAGCCCACTCTGAGTGAAGC 89
DB 26 TCTAGATGATGCTGCTCAGAGGCGCGCCAGTGTGATGATGTCAGAGATTCGCCCTTAGC 85
QY 90 TGGAGACCTCCGAGAGCCCATCATCTACTTCCCTGATGTGATGAGCCAGTCACTACCAT 149
DB 86 TGGTCCGCGGCGAGGACTCTCTCTCTCTCCCTGCTGACATATTCAGAACAGAA 145
QY 150 TCTGAGATGTTTCCCACTATCTCAGTCCGCTCCCGCAGC 190
DB 146 CATGGGGAGAGCCAGCCAGCTGCTGCTGAGCGGAGAGAGC 186
RESULT 9
AAV72224
XX ID AAV72224 standard; DNA; 333 BP.
AC AAV72224;
XX
DT 07-SEP-1999 (first entry)
XX
DE Human anti-GPIIb/IIIa antibody light chain DNA from phagemid PDG7.
XX
KW Antibody; GPIIb/IIIa; human; auto-antibody; anti-idiotypic; diagnosis;
KW blood platelet membrane protein; predisposition; prevention; treatment;
KW autoimmune thrombocytopenic purpura; AITP; fibrinogen binding; thrombi;
KW thrombocyte; cardiac infarction; pulmonary embolism; light chain; ds.
OS Homo sapiens.
XX
FH key
FT misc_feature
FT 1..60
FT location/Qualifiers
FT /tag= a
FT /note= "Framework region 1 (FR1)"
FT 61..99
FT /tag= b
FT /note= "complementarity determining region 1 (CDR1)"
FT 100..144
FT /tag= c
FT /note= "Framework region 2 (FR2)"
FT 145..165
FT /tag= d
FT /note= "complementarity determining region 2 (CDR2)"
FT 166..261
FT /tag= e
FT /note= "Framework region 3 (FR3)"

FT misc_feature 262..294
FT /tag= f
FT /note= "complementarity determining region 3 (CDR3)"
FT misc_feature 295..333
FT /tag= g
FT /note= "Framework region 4 (FR4)"
XX
PN WO9855619-A1.
XX
PD 10-DEC-1998.
XX
PF 05-JUN-1998; 98WO-EP03397.
XX
PR 08-MAY-1998; 98DE-1020663.
PR 06-JUN-1997; 97DE-1023904.
PR 12-DEC-1997; 97DE-1055227.
XX
PA (ASAT-) ASAT AG APPLIED SCI & TECHNOLOGY.
XX
PI Berchold P, Escher RFA.
XX
DR WPI; 1999-105496/09.
DR P-PSDB; AAW90279.
XX
PT Nucleic acid encoding human autoantibodies against platelet
PT glycoprotein IIb/IIIa - used for diagnosis, treatment and prevention
PT of autoimmune thrombocytopenic purpura and for modulation of
PT fibrinogen binding
XX
PS Disclosure; Page 49-50; 93pp; German.
XX
CC This invention describes novel nucleic acid fragments that encode human
CC auto-antibodies and anti-idiotypic antibodies against blood platelet
CC membrane protein, GPIIb/IIIa. The products of the invention are used
CC for diagnosis (including monitoring and determining predisposition),
CC prevention and treatment of autoimmune thrombocytopenic purpura (AITP)
CC and also for modulating binding of fibrinogen to thrombocytes
CC (particularly to dissolve thrombi and/or prevent their formation, e.g.
CC in cases of cardiac infarction or pulmonary embolism). Unlike murine
CC antibodies, human antibodies (hAb) do not induce adverse side effects
CC and persist for longer in vivo than small peptides.
XX
SQ Sequence 333 BP; 59 A; 109 C; 93 G; 72 T; 0 other;
Query Match 8.3%; Score 35.2; DB 20; Length 333;
Best Local Similarity 52.8%; Pred. No. 3.2;
Matches 76; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 10 AGCCACTGCCCCAGAGATGTCAGATCGCTGCTCAAGAGCGTCGCGAGAGAGGAC 69
DB 117 AGCCAGGCCCCCAAACTCTCATCTTTGCTAGTCAACGCGCCCTCAGGGCTCCCTGA 176
QY 70 CCAATCCTGTAGCTGAGAGCTGAGAGCTCCGAGAGCCCATTCATCTTCCCTGACTTGG 129
DB 177 CCGATTCTGTGCTCCAGTCAAGTGGGCACTCCGCTCCGTCATCGGTGCTCAATC 236
QY 130 TACAGCCACTGCTACCATTTCTG 153
DB 237 TGGGATGCTGTGACTATTACTG 260
RESULT 10
AAQ25135
XX ID AAQ25135 standard; DNA; 268 BP.
AC AAQ25135;
XX
DT 19-NOV-1992 (first entry)
XX
DE DRL.
XX
KW HLA class II; DRB; DOA; DOB; DPB; Hodgkin's disease; multiple sclerosis;
KW insulin-dependent diabetes; ss.

```

XX Synthetic.
OS
XX Key
XX misc_RNA
FT 7..21 location/Qualifiers
FT /*tag= a
FT /label= Polymorphous_region_1
FT 61..81
FT /*tag= b
FT /label= Polymorphous_region_2
XX
XX WO9208117-A.
XX 14-MAY-1992.
XX
XX 08-OCT-1991; 91WO-US07308.
XX
XX 17-OCT-1990; 90NL-0002259.
XX
XX (EURO-) EURODIAGNOSTICS BV.
XX (BIOM) APPLIED BIOSYSTEMS INC.
XX
XX Tiliannus MG:
XX
XX WPI: 1992-183800/22.
XX P-PSDB; AAR24276.
XX
XX Method and kit for determining genotype - by comparing sequences
XX of gene family members e.g. for disease detection
XX
XX Disclosure; Fig 1: 22pp; English.
XX
XX The sequence given is an allele of the human DRB locus, DRI. This
XX allele is distinguishable from other DR alleles by the sequences
XX contained within the polymorphous regions. These alleles specific
XX sequences can differ between closely related individuals, see also
XX AA025136-45 and AAR24277-86. It can be seen by haplotype analysis that
XX there are only two distinct DR alleles and the allele specific
XX sequences identify different haplotypes.
XX
XX Sequences from the HLA class II loci DQA, DQB and DRB can also be
XX analysed to produce similar information. The primer sequences used
XX in this analysis can be used determining a genotype by comparing the
XX nucleotide sequence of members of a gene system, where the sequences
XX to be compared are of a strongly conserved section of the genetic
XX material.
XX
XX This method can be used to detect genetic variations associated with
XX diseases or disease symptoms, eg. immune response defects, diseases
XX associated with the HLA system, such as Hodgkin's disease, multiple
XX sclerosis and insulin-dependent diabetes. This method may also be
XX used for tissue or cell typing, eg. to determine the degree of
XX compatibility of transplants and to determine the risk of an
XX HLA-associated disease for an animal. It allows the typing of, eg.
XX HLA class I alleles which could previously only be determined by
XX serological techniques.
XX
XX Sequence 268 BP; 58 A; 60 C; 99 G; 51 T; 0 other;
SQ
Query Match 8.0%; Score 34.2; DB 13; Length 268;
Best Local Similarity 54.3%; Pred. No. 5.7;
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 207 GGGGGCGCGCAACAGAGTCTCTGTGACCTGACGAGGAGCGCAGTTGGAACCGGG 266
DB 130 GCGGTGACGAGCTGCGGCGCGCTGATGCGAGTACGGAACGCGAGAGACCTCCTG 189
QY 267 GAGCAGAGCGAGTGGCCCGACGATCCAGTTCATGTGACCGCGGCTCATGACT 326
DB 190 GAGCAGAGCGCGCGCGGTGACACTACTGACAGACCAACTACGGGTTGTGAGAGC 249
QY 327 ATCACTG 333
DB 250 TTCACAG 256

```

```

RESULT 11
AAD31202/C
ID AAD31202 standard; DNA; 33675 BP.
XX
XX AAD31202;
AC
XX 31-MAY-2002 (first entry)
XX
XX Oryza sativa gcpe gene.
XX
XX gcpe gene; methyl-D-erythritol phosphate; MEP; dimethylallyl diphosphate;
XX transgenic plant; isoprenoid compound; tocopherol; isopentenyl diphosphate;
XX food; feed source; transfection; single nucleotide polymorphism; SNP;
XX oxidative stress tolerance; UV tolerance; transformation; rice; plant;
XX ds.
XX
XX Oryza sativa.
XX
XX Location/Qualifiers
XX 6514..11129
XX /*tag= a
XX /product= "Rice GCPE protein #1"
XX 6924..11129
XX /*tag= b
XX /product= "Rice GCPE protein #2"
XX 6514..6762
XX /*tag= c
XX /note= "this region corresponds to rice GCPE protein #1"
XX 6763..6923
XX /*tag= d
XX /note= "this region corresponds to rice GCPE protein #1"
XX /cons.splice= (5'site:YES, 3'site:NO)
XX 6924..7019
XX /*tag= e
XX /product= "Rice GCPE protein #1"
XX 7020..7162
XX /*tag= f
XX /product= "Rice GCPE protein #1"
XX 7163..7269
XX /*tag= g
XX /product= "Rice GCPE protein #1"
XX 7270..7343
XX /*tag= h
XX /product= "Rice GCPE protein #1"
XX 7344..7444
XX /*tag= i
XX /product= "Rice GCPE protein #1"
XX 7445..7524
XX /*tag= j
XX /product= "Rice GCPE protein #1"
XX 7525..7634
XX /*tag= k
XX /product= "Rice GCPE protein #1"
XX 7635..7693
XX /*tag= l
XX /product= "Rice GCPE protein #1"
XX 7694..7813
XX /*tag= m
XX /product= "Rice GCPE protein #1"
XX 7814..7922
XX /*tag= n
XX /product= "Rice GCPE protein #1"
XX 7923..8153
XX /*tag= o
XX /product= "Rice GCPE protein #1"
XX 8154..8252
XX /*tag= p
XX /product= "Rice GCPE protein #1"
XX 8253..8369
XX /*tag= q
XX /product= "Rice GCPE protein #1"
XX 8370..8514
XX /*tag= r
XX /product= "Rice GCPE protein #1"
XX 8515..8589
XX /*tag= s
XX /product= "Rice GCPE protein #1"
XX 8590..9011
XX /*tag= t
XX /product= "Rice GCPE protein #1"
XX 9012..9071
XX /*tag= u
XX /product= "Rice GCPE protein #1"
XX 9072..9162
XX /*tag= v
XX /product= "Rice GCPE protein #1"
XX 9163..9225
XX /*tag= w
XX /product= "Rice GCPE protein #1"
XX 9226..9327
XX /*tag= x
XX
XX Intron
FT

```


FT exon 9328..9472
 FT /tag= y
 FT 9473..9588
 FT intron
 FT /tag= z
 FT 9589..9730
 FT exon
 FT /tag= aa
 FT 9731..9950
 FT intron
 FT /tag= ab
 FT 9951..10028
 FT exon
 FT /tag= ac
 FT 10029..10133
 FT intron
 FT /tag= ad
 FT 10134..10293
 FT exon
 FT /tag= ae
 FT 10294..10693
 FT intron
 FT /tag= af
 FT 10694..10798
 FT exon
 FT /tag= ag
 FT 10799..11027
 FT intron
 FT /tag= ah
 FT 11028..11129
 FT exon
 FT /tag= ai

WO200212478-A2.
 14-FEB-2002.
 06-AUG-2001; 2001WO-US24335.
 07-AUG-2000; 2000US-223483P.
 (MONS) MONSANTO TECHNOLOGY LLC.
 Boronat A, Campos N, Rodriguez-concepcion M, Rohner M, Seeman M,
 PI Valentin HE, Venkatesh TV, Venkatramesh M;
 DR WPI: 2002-227151/28.
 DR P-PSDB: AAE19651, AAE19652.

gcpe nucleic acid which is an essential gene of the methyl-D-erythritol
 phosphate pathway, encoding a fully defined GCPE protein which is
 useful for increasing levels of tocopherol substrates in plants -

Claim 3; Page 97-117; 155pp: English.

The invention relates to gcpe nucleic acid molecule, an essential gene
 of methyl-D-erythritol phosphate (MEP) pathway) that encodes rice,
 CC Arabidopsis thaliana or Escherichia coli GCPE protein. gcpe is useful
 CC for producing a transgenic plant such as Brassica campestris, B. napus,
 CC canola, castor bean, coconut, cotton, crambe, linseed, maize, mustard,
 CC oil palm, peanut, rapeseed, rice, safflower, sesame, soybean, sunflower,
 CC or wheat with an increased isoprenoid (tocopherol) compound level. The
 CC expression of GCPE protein in organisms increases the level of
 CC tocopherol substrate such as isopentenyl diphosphate and dimethylallyl
 CC diphosphate biosynthesis. Transgenic organisms overexpressing GCPE
 CC protein can nutritionally enhance food and feed sources. Overexpression
 CC of GCPE protein in transgenic plant may provide tolerance to stresses
 CC e.g., oxidative stress tolerance such as to oxygen or ozone, UV
 CC tolerance, etc. gcpe may be used to obtain nucleic acid molecules from
 CC the same species, and to obtain nucleic acid homologues. gcpe is also
 CC used as or primers. The recombinant vectors are used in plant
 CC transformation or transfection. gcpe an also act as markers capable of
 CC detecting polymorphisms such as single nucleotide polymorphisms (SNPs).
 CC gcpe is also used to determine the level or pattern of expression of
 CC the protein. The present sequence is Oryza sativa gcpe gene.

Sequence 33675 BP; 9475 A; 7073 C; 7297 G; 9801 T; 29 other;
 Query Match 7.9%; Score 33.6; DB 24; Length 33675;
 Best Local Similarity 55.0%; Pred. No. 37;
 Matches 66; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

229 CTTGTGACCTGACGAGCGCGAGTGTGAACCCCGGAGAGACGACGAGTGGCCACAG 288

DB 12658 CTCGGGAGGAGCCCGGTCGTCTGTATCATCGCGCGGAAGGGCGGTGGAGAG 12539
 OY 289 GTACCATGCGATTTCATGTCACCGCGGTGTATGACTGCAACATCTACATCT 348
 DB 12598 GCATCAGCGGACATCTACACTTCGGGAGCTCATCTCGGACACCGCAACTCTTCGCTT 12539

RESULT 12
 AAS01024
 ID AAS01024 standard; cDNA; 955 BP.
 AC AAS01024;
 XX
 DT 29-MAY-2001 (first entry)
 XX
 DE Sugarcane plant gene promoter cDNA isolated from clone c512.
 XX
 KW Sugarcane promoter region; monocotyledonous plant; stem tissue;
 KW insecticide; herbicide; disease resistance; improved food content;
 KW beta-glucuronidase; GUS; starch biosynthesis; fatty acid biosynthesis;
 KW ADP-glucose pyrophosphorylase; sucrose metabolism; clone c512; ss.
 XX
 OS Saccharum sp.
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..122
 FT /tag= a
 FT CDS 123..572
 FT /tag= b
 FT /note= "Amino acid sequence deduced from the
 FT 3'UTR 573..924
 FT /tag= c
 FT polyA_site 925
 FT /tag= d
 PN WO200118211-A1.
 XX
 PD 15-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-AU01033.
 XX
 PR 02-SEP-1999; 99AU-0002625.
 XX
 PA (UYOU) UNIV QUEENSLAND.
 XX
 PI Potler B, Birch RG;
 XX
 DR WPI: 2001-218560/22.
 DR P-PSDB: AAU00451.

New sugarcane plant promoters for directing expression of heterologous
 nucleic acids in a constitutive or tissue-specific manner in
 monocotyledonous plants -

Claim 1; Fig 15; 107pp; English.

The present sequence for sugarcane plant promoter cDNA isolated from
 clone c512 is 1 of 11 promoter regions of a transcribable DNA sequence
 isolated from various sugarcane cDNA clones (AAS01021-AAS01031).
 CC Clones c51, c511 and c512 are homologous cDNA sequences. Also
 CC described are 4 promoter regions of specific transcribed DNA sequences
 CC (AAS01032-AAS01035). The nucleic acids are useful for producing
 CC transgenic plants, having an altered phenotype and for driving expression
 CC of a foreign or endogenous DNA sequence, which encode agronomic
 CC properties including insecticide, herbicide, disease resistance,
 CC stress tolerance and improved food content, or increased yields. The
 CC foreign or endogenous DNA sequence may comprise a region transcribed
 CC into an antisense RNA or ribozyme that modulates the expression of a
 CC corresponding target gene, or it may encode beta-glucuronidase (GUS),
 CC luciferase, neomycin phosphotransferase, a product conferring herbicide
 CC tolerance, a product affecting starch biosynthesis or modification,

PT New sugarcane plant promoters for directing expression of heterologous nucleic acids in a constitutive or tissue-specific manner in

[illegible]

PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
XX WPI: 2000-072177/06.
DR
XX Novel polynucleotides useful for the treatment of various conditions
PT including wounds and cancer -
XX
PS Claim 1: Page 69-70; 235pp; English.
XX
CC The invention relates to novel nucleic acid sequences derived from rat
CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
CC cells. Polypeptides of the invention may be used to treat inflammation,
CC cancer and neurological diseases. The proteins may be used to stimulate
CC the growth and motility of keratinocytes, to inhibit the growth of
CC cancer cells, to modulate angiogenesis and tumour vascularisation, to
CC modulate skin inflammation, to modulate epithelial cell growth and to
CC inhibit binding of HIV-1 to leukocytes. The invention may also be used
CC to treat growth and developmental defects, skin wounds and hair follicle
CC disorders. Sequences AA261606-261832 represent cDNA sequences derived
CC from several mouse, rat or human skin cell types. Sequences
CC AA261606-261649, AA261725-261765, AA261802-261811 and AA261826 encode
CC proteins with an N-terminal signal sequence, indicating that the proteins
CC are secreted. Sequences AA261650-261668, AA261766-261780, AA261812-261817
CC and AA261827-261829 encode proteins with one or more putative
CC transmembrane domains.
XX
SQ Sequence 1057 BP; 223 A; 315 C; 295 G; 224 T; 0 other:

Query Match 7.8%; Score 33; DB 21; Length 1057;
Best Local Similarity 51.0%; Pred. No. 19;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 53 TCCGAGGAGAGGAGACCCATCTGTAGCTGGAAGCTGGAGCCTCCGAGGCCCATCC 112
DB 353 TCCGACACGCTGGGATTCGATCTGCAGCGGAGAGAGATGCCATTACAGGCTCTTCC 412
QY 113 ATACTTCCCTGACTGTGTACAGCCACTGTACCACTTTCTGAGATGTTTCCCAATATC 172
DB 413 CTAAGCCCTGGACAGATGATCTCCAAAGTACTCTCTCGGGAGCTACACCTGTATCA 472
QY 173 CACTGGGCTCCCGGAGCCCGCAGTTTGGAGGC 205
DB 473 CGCAAGGCTTTTGGAGAGACCGCATACTGGGGC 505

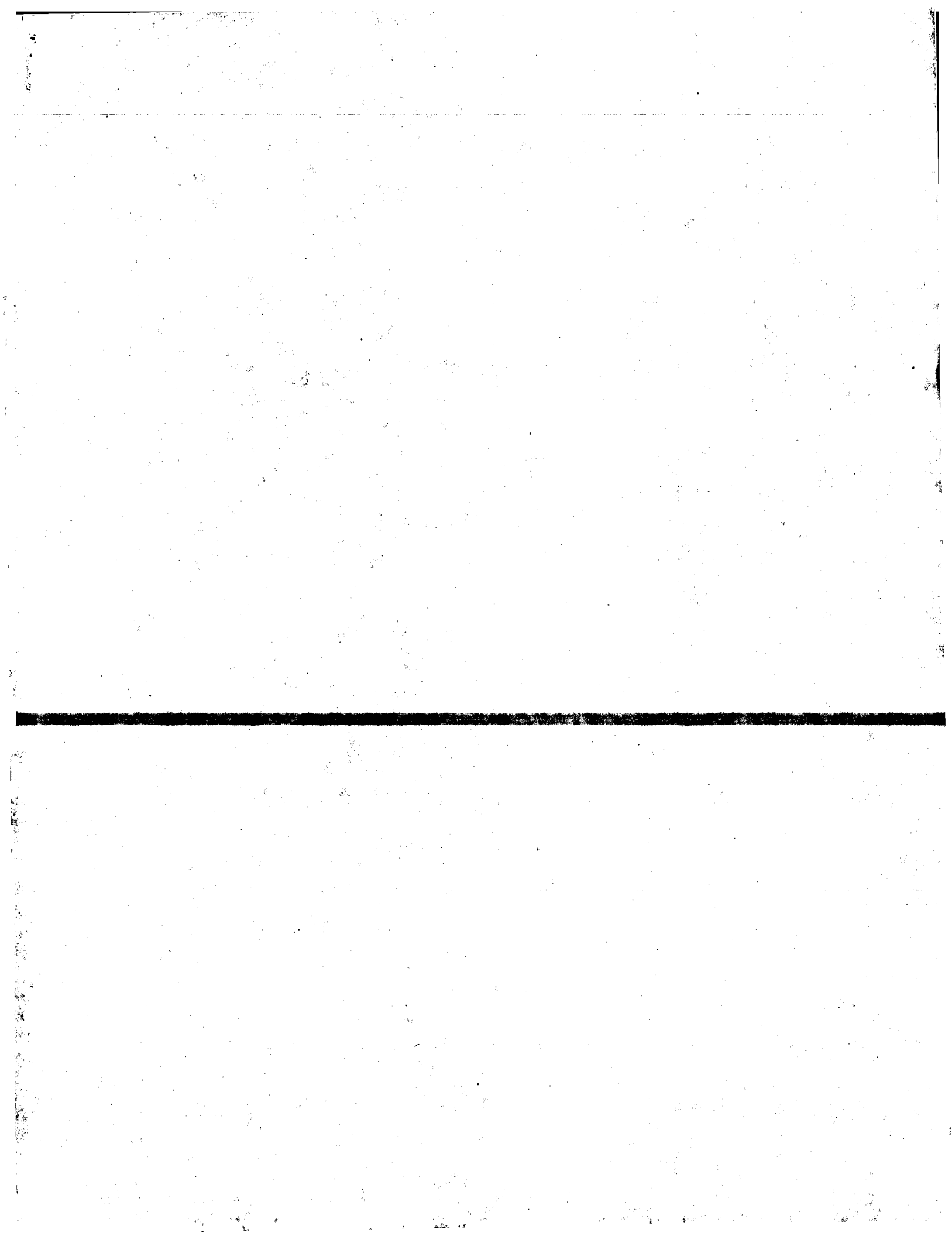
RESULT 15
AAC99556
ID AAC99556 standard; cDNA; 1057 BP.
XX
AC AAC99556;
XX
DT 08-MAR-2001 (first entry)
XX
DE Skin cell cDNA, SEQ ID NO: 18.
XX
KM Rat; skin cell; cytosolic; anti-inflammatory; anti-HIV;
KM neurotropic; neuroprotective; vulnery; immunomodulatory; vaccine;
KM keratinocyte growth stimulation; cancer; angiogenesis inhibition;
KM inflammation; neurological disease; ss.
XX
OS Rattus sp.
XX
FN WO200069884-A2.
XX
PD 23-NOV-2000.
XX
PF 15-MAY-2000; 2000WO-N200075.
XX
PR 14-MAY-1999; 99US-0312283.
XX

PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD, Murison JG;
XX WPI: 2001-007495/01.
DR P-PSDB: AAB55898.
XX
XX New isolated polynucleotide used in the identification of genetic
PT disorders and encoding polypeptides used for treating inflammatory
PT disease, cancer and neurological diseases -
XX
PS Claim 1: Page 81; 352pp; English.
XX
CC The present polynucleotide encodes a polypeptide which is expressed in
CC mammalian skin cells. The polypeptide is useful for stimulating
CC keratinocyte growth and motility, inhibiting the growth of cancer cells,
CC modulating angiogenesis, inhibiting angiogenesis and vascularisation of
CC tumours, modulating skin inflammation, stimulating the growth of
CC epithelial cells, inhibiting the binding of human immunodeficiency virus
CC (HIV)-1 to leukocytes, and treating inflammatory disease, cancer and
CC neurological diseases. The polynucleotide can be used as a marker, in
CC the identification of genetic disorders, and for the design of
CC oligonucleotides for examining expression patterns.
XX
SQ Sequence 1057 BP; 223 A; 315 C; 295 G; 224 T; 0 other:

Query Match 7.8%; Score 33; DB 22; Length 1057;
Best Local Similarity 51.0%; Pred. No. 19;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 53 TCCGAGGAGAGGAGACCCATCTGTAGCTGGAAGCTGGAGCCTCCGAGGCCCATCC 112
DB 353 TCCGACACGCTGGGATTCGATCTGCAGCGGAGAGAGATGCCATTACAGGCTCTTCC 412
QY 113 ATACTTCCCTGACTGTGTACAGCCACTGTACCACTTTCTGAGATGTTTCCCAATATC 172
DB 413 CTAAGCCCTGGACAGATGATCTCCAAAGTACTCTCTCGGGAGCTACACCTGTATCA 472
QY 173 CACTGGGCTCCCGGAGCCCGCAGTTTGGAGGC 205
DB 473 CGCAAGGCTTTTGGAGAGACCGCATACTGGGGC 505

Search completed: April 15, 2003, 15:16:58
Job time : 95.9741 secs



GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:09:45 ; Search time 490.898 Seconds

(without alignments)
14021.414 Million cell updates/sec

Title: US-09-917-372-2_COPY_853_1277

Perfect score: 425
Sequence: 1 atccattagacgcactgcgc.....aacctccatccattccc 425

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	393	92.5	814	9 AU118203	AU118203 AU118203
2	393	92.5	868	10 BE275240	BE275240 601122062
3	393	92.5	895	14 BQ957322	BQ957322 AGENCOURT
4	393	92.5	920	12 BG036632	BG036632 602326834
5	393	92.5	954	13 B1821009	B1821009 603035664
6	393	92.5	968	14 BQ898015	BQ898015 AGENCOURT

7	388	91.3	427	10 AM630662	AM630662 hb85h07.y
8	381	89.6	640	12 BG491235	BG491235 602535283
9	381	89.6	714	13 B1829057	B1829057 603079147
10	381	89.6	900	13 B1769498	B1769498 603059047
11	381	89.6	901	12 BE740143	BE740143 601355085
12	380.4	89.5	954	14 BQ930450	BQ930450 AGENCOURT
13	370	87.1	700	12 BG576875	BG576875 602599233
14	355	83.5	749	12 BE789430	BE789430 601482267
15	346.6	81.6	579	10 BE207591	BE207591 bbb6f10.y
16	344	80.9	725	12 BG331666	BG331666 602432673
17	343.8	80.9	606	13 B1905922	B1905922 603063085
18	342	80.5	598	12 BG403820	BG403820 602419660
19	326	76.7	728	13 B1765703	B1765703 603046670
20	322	75.8	801	12 BF792867	BF792867 602253351
21	316	74.4	458	14 BM766807	BM766807 K-EST0048
22	316	74.4	501	14 BM766833	BM766833 K-EST0048
23	316	74.4	537	14 BM766352	BM766352 K-EST0048
24	316	74.4	572	14 BM767287	BM767287 K-EST0049
25	316	74.4	666	14 BM767318	BM767318 K-EST0049
26	314.4	74.0	634	14 BM746911	BM746911 K-EST0021
27	311	73.2	598	14 BM723050	BM723050 UT-E-EO1-
28	301.8	71.0	360	14 BM766382	BM766382 K-EST0048
29	293.4	69.0	366	9 AA099514	AA099514 z177h03.r
30	282	66.4	598	14 BQ305454	BQ305454 MRO-BT200
31	275	64.7	506	14 BM855039	BM855039 K-EST0137
32	274.4	64.6	276	14 BM784961	BM784961 K-EST0063
33	274.4	64.5	808	13 B1818411	B1818411 603032875
34	273.4	64.3	466	14 BQ082995	BQ082995 K-EST0144
35	269.8	63.5	426	14 BM856797	BM856797 K-EST0140
36	263.4	62.0	416	14 BQ083004	BQ083004 K-EST0144
37	261.8	61.6	418	14 BM855410	BM855410 K-EST0138
38	253.8	59.7	281	9 AA476989	AA476989 z031d09.r
39	251.2	59.1	289	9 AA479617	AA479617 z030303075
40	249.2	58.1	942	13 B1818552	B1818552 603033075
41	247	58.1	346	14 H59746	H59746 yr38e10.r1
42	246.6	58.0	721	9 A1336202	A1336202 q144e06.x
43	245.4	57.7	530	12 BG610053	BG610053 324524 MA
44	244	57.4	352	10 AM605918	AM605918 R01-HT025
45	236	55.5	706	12 BG698325	BG698325 602658595

ALIGNMENTS

RESULT 1	AU118203	814 bp	mRNA	linear	EST 01-AUG-2002
LOCUS	AU118203	HEMBA1	Homo sapiens	cdna	clone HEMBA1003089 5', mRNA
DEFINITION	AU118203	sequence.			
ACCESSION	AU118203				
VERSION	AU118203.1	GI:10933231			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 814)				
AUTHORS	Yamanoto,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Isogai,T.				
TITLE	HRI human cDNA project				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.				

FEATURES	source	Location/Qualifiers
		1.. 814
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="HEMBA1003089"
		/clone_id="HEMBA1"
		/tissue_type="whole embryo, mainly head"
		/dev_stage="embryo, 10 weeks"
		/note="Vector: pME18SFL3"
BASE COUNT	173 a 255 c 232 g 151 t	3 others
ORIGIN		
Query Match	92.5%: Score 393; DB 9; Length 814;	
Best Local Similarity	100.0%: Pred. No. 9.7e-87;	
Matches 393; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	33 GGATGCTGCTCAAGAGGCGTCCGAGGAGAGGAGACCCATCTGTAGCTGGAAGCTGG	92
DB	98 GGATGGCTCTCTAAGAGGCGTCCGAGGAGAGGAGACCCATCTGTAGCTGGAAGCTGG	157
QY	93 GAGCCTCGAAGGCGCCATTCATCTTCCCTGACTTGTACAGACCACTGTACCACTTCT	152
DB	158 GAGCCTCGAAGGCGCCATTCATCTTCCCTGACTTGTACAGACCACTGTACCACTTCT	217
QY	153 GGAGATGTTTCCCAAGTATCCACTGGGCTCCCCGAGCCCGAGTTTGGAGCAGGGTG	212
DB	218 GGAGATGTTTCCCAAGTATCCACTGGGCTCCCCGAGCCCGAGTTTGGAGCAGGGTG	277
QY	213 CGCAACAGCAGAGTCTGTGACCTGACAGGAGGCGCAGTTGGAACCCGGGGAGCAG	272
DB	278 CGCAACAGCAGAGTCTGTGACCTGACAGGAGGCGCAGTTGGAACCCGGGGAGCAG	337
QY	273 AGCCAGGTGGCCCAAGTACCAATGGCATTCATGTCACCGGGGGTCTATGACTATCACT	332
DB	338 AGCCAGGTGGCCCAAGTACCAATGGCATTCATGTCACCGGGGGTCTATGACTATCACT	397
QY	333 GGCACATCTCATCTACAAATGACACAGTACGAGGGGGGAGCACCCGGGTCCTGGAACCTC	392
DB	398 GGCACATCTCATCTACAAATGACACAGTACGAGGGGGGAGCACCCGGGTCCTGGAACCTC	457
QY	393 CCAGCTACCCCGAAGCTTCATACCCCAATTTCC	425
DB	458 CCAGCTACCCCGAAGCTTCATACCCCAATTTCC	490
RESULT 2		
BE275240	868 bp	mRNA
LOCUS	60112206221 NIH_MGC_20	Homo sapiens cDNA clone IMAGE:346311 5',
DEFINITION	mRNA sequence.	
ACCESSION	BE275240	
VERSION	BE275240.1	GI:9150193
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 868)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999).	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-remail.nih.gov Tissue Procurement: ATCC/DC/DP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov Plate: LCM135 row: k column: 16 High quality sequence stop: 746. Location/Qualifiers 1.. 868	

Query Match	Best Local Similarity	Score 393;	DB 10;	Length 868;
Matches 393;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps				
33	GGATGCTGCTCTAAGAGGGGTCCGAGGAGAGGGAGCCCAATCTGTAGCTGGAAGCTGG	92.5%	DB 10;	Length 868;
138	GGATGCTGCTCTAAGAGGGGTCCGAGGAGAGGGAGCCCAATCTGTAGCTGGAAGCTGG	100.0%	Pred. No. 9.9e-87;	
93	GAGCCTCGGAAGGGCCCATTCATCTTCCCTGACTTGGTACAGCCACTGTACCCATTCT	0;	Indels	0;
198	GAGCCTCGGAAGGGCCCATTCATCTTCCCTGACTTGGTACAGCCACTGTACCCATTCT	0;	Gaps	0;
153	GGAGATGTTTCCCGACGTATCCACTGGGCTCCCGGAGCCCGAGCTTTGGAGGAGGGGTG			
258	GGAGATGTTTCCCGACGTATCCACTGGGCTCCCGGAGCCCGAGCTTTGGAGGAGGGGTG			
213	CGGCAACAGCAGAGTCTCTGTGACCTGACACAGGAGCGCATTTGGAACCCCGGGAGCAG			
318	CGGCAACAGCAGAGTCTCTGTGACCTGACACAGGAGCGCATTTGGAACCCCGGGAGCAG			
273	AGCCAGGTGGCCCGACGTACCAATGGCATTTGTCACCGGGGGTCTATGACTATCACT			
378	AGCCAGGTGGCCCGACGTACCAATGGCATTTGTCACCGGGGGTCTATGACTATCACT			
333	GGCACAATCTACATCTACAAATGAGACAGTACTGGGGGGGACCAACCGGGTCTGGAAGCTGC			
438	GGCACAATCTACATCTACAAATGAGACAGTACTGGGGGGGACCAACCGGGTCTGGAAGCTGC			
393	CCAGCTACCCCGGACGCTTCATACCCCATTTGCC	425		
498	CCAGCTACCCCGGACGCTTCATACCCCATTTGCC	530		

RESULT 3

LOCUS B0957322

DEFINITION B0957322

ACCESSION B0957322

VERSION B0957322.1

KEYWORDS GI:22372800

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs@remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.W.A.G.E. Consortium (LMD)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

for organism="Homo sapiens"

Ad_xref="taxon:9606"

clone="IMAGE:3346311"

clone_lib="NIH_MGC_20"

tissue_type="melanocytic melanoma"

lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2579 row: g column: 17
High quality sequence stop: 645.

FEATURES

source
1. 895
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6384424"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; CDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 190 a 291 c 252 g 159 t 3 others
ORIGIN

Query Match 92.5%; Score 393; DB 14; Length 895;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 GGATCGCTGCTCAGAGGGGCTCCGAGGAGGAGGCCCAATCCTGAGTGAAGCTGG 92
132 GGATCGCTGCTCAGAGGGGCTCCGAGGAGGAGGCCCAATCCTGAGTGAAGCTGG 191
93 GAGCCTCCGAGAGGCCCATCTACTTCCCTGAGTGAAGGAGGAGGAGGAGGAGG 152
192 GAGCCTCCGAGAGGCCCATCTACTTCCCTGAGTGAAGGAGGAGGAGGAGGAGG 251
153 GAGATGTTTCCCGATATTCAGTGGGCTCCCGAGGCCCATTTTGGAGGAGGGTG 212
252 GGAGATGTTTCCCGATATTCAGTGGGCTCCCGAGGCCCATTTTGGAGGAGGGTG 311
213 CCGCAACAGACAGAGTCTCTGACCTGACAGGAGGAGGAGGAGGAGGAGGAGGAG 272
312 CCGCAACAGACAGAGTCTCTGACCTGACAGGAGGAGGAGGAGGAGGAGGAGGAG 371
273 AGCAGAGTGGCCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 332
372 AGCAGAGTGGCCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
333 GGCACATCTACATCTACATGAGACAGTACTGGGGGAGACACCGGGTCTGGAGACTC 392
432 GGCACATCTACATCTACATGAGACAGTACTGGGGGAGACACCGGGTCTGGAGACTC 491
393 CCAGCTACCCCGAGACCTCCATACCCCATTTCC 425
492 CCAGCTACCCCGAGACCTCCATACCCCATTTCC 524
Db

RESULT 4
LOCUS BG036632 920 bp mRNA linear EST 24-JAN-2001

DEFINITION 603226834F1 NIH_MGC_91 Homo sapiens CDNA clone IMAGE:442825 5',
mRNA sequence.

ACCESSION BG036632
VERSION BG036632.1 GI:12432013
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 920)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov
Tissue Procurement: DCM/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10178 row: c column: 10
High quality sequence stop: 693.

FEATURES

source
1. 920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:442825"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-Sport6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 187 a 304 c 270 g 159 t
ORIGIN

Query Match 92.5%; Score 393; DB 12; Length 920;
Best Local Similarity 100.0%; Pred. No. 1e-86;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 GGATCGCTGCTCAGAGGGGCTCCGAGGAGGAGGCCCAATCCTGAGTGAAGCTGG 92
52 GGATCGCTGCTCAGAGGGGCTCCGAGGAGGAGGCCCAATCCTGAGTGAAGCTGG 111
93 GAGCCTCCGAGAGGCCCATCTACTTCCCTGAGTGAAGGAGGAGGAGGAGGAGGAG 152
112 GAGCCTCCGAGAGGCCCATCTACTTCCCTGAGTGAAGGAGGAGGAGGAGGAGGAG 171
153 GAGATGTTTCCCGATATTCAGTGGGCTCCCGAGGCCCATTTTGGAGGAGGGTG 212
172 GGAGATGTTTCCCGATATTCAGTGGGCTCCCGAGGCCCATTTTGGAGGAGGGTG 231
213 CCGCAACAGACAGAGTCTCTGACCTGACAGGAGGAGGAGGAGGAGGAGGAGGAG 272
232 CCGCAACAGACAGAGTCTCTGACCTGACAGGAGGAGGAGGAGGAGGAGGAGGAG 291
273 AGCAGAGTGGCCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 332
292 AGCAGAGTGGCCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 351
333 GGCACATCTACATCTACATGAGACAGTACTGGGGGAGACACCGGGTCTGGAGACTC 392
352 GGCACATCTACATCTACATGAGACAGTACTGGGGGAGACACCGGGTCTGGAGACTC 411
393 CCAGCTACCCCGAGACCTCCATACCCCATTTCC 425
412 CCAGCTACCCCGAGACCTCCATACCCCATTTCC 444
Db

RESULT 5
LOCUS B1821009 954 bp mRNA linear EST 04-OCT-2001

DEFINITION 603035664F1 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5176935 5',
mRNA sequence.

ACCESSION B1821009
VERSION B1821009.1 GI:15932559
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 954)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.

Db	626	COAGTACCCCGAACCCTCCATPACCCCACTTCC	658
RESULT 7		427 bp	mRNA
LOCUS	AM630662		linear
DEFINITION	h855h07.y1 NCI_CGAP_GUI Homo sapiens cDNA clone IMAGE:2969629 5' similar to db:L04270 TUMOR NECROSIS FACTOR RECEPTOR 2 RELATED PROTEIN PRECURSOR (HUMAN); contains element MSRI repetitive element ;, mRNA sequence.		EST 31-MAR-2000
ACCESSION	AM630662		
VERSION	AM630662.1	GI:7377452	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 427)		
TITLE	NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap .		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgaps-remail.nih.gov Tissue Procurement: Chris Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/html/lresources.shtml Seq primer: -40RP from Glbco High quality sequence stop: 340. Location/Qualifiers 1..427 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2969629" /clone_lib="NCI CGAP GUI" /tissue_type="2 pooled high-grade transitional cell tumors" /lab_host="DH10B" /note="Organ: genitourinary tract; Vector: pCMV-Sport6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: oligo dT. Library constructed by Life Technologies."		
BASE COUNT	89 a 145 c 118 g 75 t		
ORIGIN			
Query Match	91.3%; Score 388; DB 10; Length 427;		
Best Local Similarity	100.0%; Pred. No. 1,3e-85;		
Matches 388;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Y	38 GCTGCTCAAGAGGCGTCCGCAAGAGAGAGACCAATCCTGTAGCTGAGAGCTGGAGCC	97	
Db	1 GCTGCTCAAGAGGCGTCCGCAAGAGAGAGAGACCAATCCTGTAGCTGAGAGCTGGAGCC	60	
Y	98 TCCGAAGGCCCATCCATCTCTCCGACTGTGGTACAGACACGTGACCAATTTCTGGAGA	157	
Db	61 TCCGAAGGCCCATCCATCTCTCCGACTGTGGTACAGACACGTGACCAATTTCTGGAGA	120	
Y	158 TGTTCCTCCAGTATCACTGTGGGCTCCCGCAAGCCCAATTTGGAGAGAGGGTCCGCA	217	
Db	121 TGTTCCTCCAGTATCACTGTGGGCTCCCGCAAGCCCAATTTGGAGAGAGGGTCCGCA	180	
Y	218 ACAGAGAGTCTCTGTGACCTGACAGGAGAGCCGACGATTGGAAACCCGGGAGACAGCCA	277	
Db	181 ACAGAGAGTCTCTGTGACCTGACAGGAGAGCCGACGATTGGAAACCCGGGAGACAGCCA	240	
Y	278 GGTGCCCCACGGTACCAATGGCATTCATGATCAACGGGGGCTATGACTATACATGCGCA	337	
Db	241 GGTGCCCCACGGTACCAATGGCATTCATGATCAACGGGGGCTATGACTATACATGCGCA	300	

QY	338	CATCAGACATCATCATGAGACAGTCTGAGGGGAGACACGGGCTCTGGAGACCTCCAGC	397
Db	301	CATCATCATCATCATGAGACAGTCTGAGGGGAGACACGGGCTCTGGAGACCTCCAGC	360
QY	398	TACCCCGAACCTCCATACCCCATTC	425
Db	361	TACCCCGAACCTCCATACCCCATTC	388
RESULT 8			
LOCUS	BC491235		
DEFINITION	BC491235.	640 bp	linear EST 27-MAR-2001
ACCESSION	602535283	1	Homo sapiens cDNA clone IMAGE:4654076 5', mRNA sequence.
VERSION	BC491235		
KEYWORDS	BC491235.1	GI:13452747	
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 640)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
FEATURES	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: DCTD/DP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LILCM141 row: e column: 21 High quality sequence start: 199 High quality sequence stop: 640.		
SOURCE	Location/Qualifiers		
	1..640		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4654076"		
	/clone_lib="NIH-MGC 41"		
	/tissue_type="amelanotic melanoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGAGAG(5). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."		
BASE COUNT	134 a	205 c	159 g
ORIGIN	142 t		
Query Match	89.6%	Score 381;	DB 12; Length 640;
Best Local Similarity	100.0%	Pred. No. 8.2e-84;	
Matches 381;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	33	GGATGCTGCTCAAGAGGGCTCCGAGAGAGAGGAGACCAATCTGTAGCTGGAAGCTGG	92
Db	260	GGATGCTGCTCAAGAGGGCTCCGAGAGAGAGGAGACCAATCTGTAGCTGGAAGCTGG	319
QY	93	GAGCTCGAGAGCCCATCATCATCTTCCCTGACTTGCTAGACCACTGCTACCATTTCT	152
Db	320	GAGCTCGAGAGCCCATCATCATCTTCCCTGACTTGCTAGACCACTGCTACCATTTCT	379
QY	153	GAGAGATGTTTCCCGAGTATCCACTGGGCTCCCGAGCCCGAGTTTGGAGGACAGGGG	212
Db	380	GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGAGCCCGAGTTTGGAGGACAGGGG	439

QY 213 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGCCGAGTTGGACCCGGGAGCAG 272
 |||||||
 Db 440 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGCCGAGTTGGACCCGGGAGCAG 499
 |||||||
 QY 273 AGCCAGGTGGCCACGAGTACCAATGATGATGACCGGGGCTATGACTATCACT 332
 |||||||
 Db 500 AGCCAGGTGGCCACGAGTACCAATGATGATGACCGGGGCTATGACTATCACT 559
 |||||||
 QY 333 GGCACATCTACATCTACATGACGACGATGAGGGGAGCACCAGGCTCTGAGAGACTC 392
 |||||||
 Db 560 GGCACATCTACATCTACATGACGACGATGAGGGGAGCACCAGGCTCTGAGAGACTC 619
 |||||||
 QY 393 CCAGCTACCCCGAAGCTTCA 413
 |||||||
 Db 620 CCAGCTACCCCGAAGCTTCA 640
 |||||||

RESULT 9
 B1829057 714 bp mRNA linear EST 04-OCT-2001
 LOCUS 603079147F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5171122 5',
 DEFINITION mRNA sequence.
 ACCESSION B1829057
 VERSION B1829057.1 GI:15940607
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 714)
 AUTHORS NIH-MGC http://mhc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLAM1425 row: m column: 11
 High quality sequence stop: 711.

FEATURES
 source
 Location/Qualifiers
 1..714
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5171122"
 /clone_lib="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: Brain; Vector: pCMV-SPORT6; Site:1: NotI;
 Site:2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC Library."

BASE COUNT 152 a 231 c 202 g 129 t

ORIGIN
 Query Match 89.6%; Score 381; DB 13; Length 714;
 Best Local Similarity 99.7%; Pred. No. 8.5e-84;
 Matches 392; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 33 GATGCGTCTCTCAAGAGGCGTCCGAGGAGGAGGACCAATCTGAGTGAAGCTGG 92
 |||||||
 Db 81 GATGCGTCTCTCAAGAGGCGTCCGAGGAGGAGGACCAATCTGAGTGAAGCTGG 140
 |||||||
 QY 93 GAGCTCCGAAAGGCCCATCATCTTCCCTGATCTGTACAGCCACTGCTACCATTTCT 152
 |||||||

Db 141 GAGCTCCGAAAGGCCCATCATCTTCCCTGATCTGTACAGCCACTGATACCATTTCT 200
 |||||||
 QY 153 GAGATGTTTCCCAAGATATACATGAGGCTCCCGACCCAGTTTGGAGGAGGAGG 212
 |||||||
 Db 201 GAGATGTTTCCCAAGATATACATGAGGCTCCCGACCCAGTTTGGAGGAGGAGG 260
 |||||||
 QY 213 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGCCGAGTTGGACCCGGGAGCAG 272
 |||||||
 Db 261 CCGCAACAGCAGAGTCTCTGACCTGACAGGAGCCGAGTTGGACCCGGGAGCAG 319
 |||||||
 QY 273 AGCCAGGTGGCCACGAGTACCAATGATGATGACCGGGGCTATGACTATCACT 332
 |||||||
 Db 320 AGCCAGGTGGCCACGAGTACCAATGATGATGACCGGGGCTATGACTATCACT 379
 |||||||
 QY 333 GGCACATCTACATCTACATGACGACGATGAGGGGAGCACCAGGCTCTGAGAGACTC 392
 |||||||
 Db 380 GGCACATCTACATCTACATGACGACGATGAGGGGAGCACCAGGCTCTGAGAGACTC 439
 |||||||

RESULT 10
 B1769498 900 bp mRNA linear EST 25-SEP-2001
 LOCUS 603059047F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5208394 5',
 DEFINITION mRNA sequence.
 ACCESSION B1769498
 VERSION B1769498.1 GI:15761076
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 900)
 AUTHORS NIH-MGC http://mhc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLAM1522 row: n column: 11
 High quality sequence stop: 854.

FEATURES
 source
 Location/Qualifiers
 1..900
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5208394"
 /clone_lib="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
 Site:1: NotI; Site:2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lungs, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH_MGC Library."

BASE COUNT 199 a 298 c 246 g 157 t

ORIGIN
 Query Match 89.6%; Score 381; DB 13; Length 900;
 Best Local Similarity 99.7%; Pred. No. 9.2e-84;

Matches 392; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 33 GGATGCTGCTCTCAAGAGGGTCCGAGAGGAGGAGCCCAATCCTGTACTGGAAGCTGG 92
|||||
Db 269 GGATGCTGCTCTCAAGAGGGTCCGAGAGGAGGAGCCCAATCCTGTACTGGAAGCTGG 328
QY 93 GAGCTCCGAAAGGCCCATCATCTACTTCCCTGACTTGGTACAGCCACTGTACCCATTCT 152
|||||
Db 329 GAGCTCCGAAAGGCCCATCATCTACTTCCCTGACTTGGTACAGCCACTGTACCCATTCT 388
QY 153 GAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGATTGGAGGAGGGGTG 212
|||||
Db 389 GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGATTGGAGGAGGGGTG 448
QY 213 CGGCAACAGCAGAGTCTGTGACCTGACAGGAGGAGCCCAATTGGAACCCGGAGGAG 272
|||||
Db 449 CGGCAACAGCAGAGTCTGTGACCTGACAGGAGGAGCCCAATTGGAACCCGGAGGAG 507
QY 273 AGCCAGTGGCCCGACGATGCAATGATGTCACGGGGGTCTATGACTATCACT 332
|||||
Db 508 AGCCAGTGGCCCGACGATGCAATGATGTCACGGGGGTCTATGACTATCACT 567
QY 333 GGCACATCTACATCTACAAATGAGACCACTGAGGGGGAGCACCGGGGTCTGAGACCTC 392
|||||
Db 568 GGCACATCTACATCTACAAATGAGACCACTGAGGGGGAGCACCGGGGTCTGAGACCTC 627
QY 393 CGAGTACCCCGGACCTCCATACCCCATTTCC 425
|||||
Db 628 CCAGTACCCCGGACCTCCATACCCCATTTCC 660

RESULT 11
BE740143 901 bp mRNA linear EST 15-SEP-2000
LOCUS BE740143 601595085F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949148 5',
DEFINITION mRNA sequence.
ACCESSION BE740143
VERSION BE740143.1 GI:10154135
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 901)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LINC813 row: 1 column: 21
High quality sequence stop: 745.
Location/Qualifiers
1..901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949148"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 181 a 297 c 267 g 156 t
ORIGIN

Query Match 89.6%; Score 381; DB 12; Length 901;
Best Local Similarity 99.7%; Pred. No. 9.2e-84;
Matches 392; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 33 GGATGCTGCTCTCAAGAGGGTCCGAGAGGAGGAGCCCAATCCTGTACTGGAAGCTGG 92
|||||
Db 55 GGATGCTGCTCTCAAGAGGGTCCGAGAGGAGGAGCCCAATCCTGTACTGGAAGCTGG 113
QY 93 GAGCTCCGAAAGGCCCATCATCTACTTCCCTGACTTGGTACAGCCACTGTACCCATTCT 152
|||||
Db 114 GAGCTCCGAAAGGCCCATCATCTACTTCCCTGACTTGGTACAGCCACTGTACCCATTCT 173
QY 153 GAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGATTGGAGGAGGGGTG 212
|||||
Db 174 GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGATTGGAGGAGGGGTG 233
QY 213 CGGCAACAGCAGAGTCTGTGACCTGACAGGAGGAGCCCAATTGGAACCCGGAGGAG 272
|||||
Db 234 CGGCAACAGCAGAGTCTGTGACCTGACAGGAGGAGCCCAATTGGAACCCGGAGGAG 293
QY 273 AGCCAGTGGCCCGACGATGCAATGATGTCACGGGGGTCTATGACTATCACT 332
|||||
Db 294 AGCCAGTGGCCCGACGATGCAATGATGTCACGGGGGTCTATGACTATCACT 353
QY 333 GGCACATCTACATCTACAAATGAGACCACTGAGGGGGAGCACCGGGGTCTGAGACCTC 392
|||||
Db 354 GGCACATCTACATCTACAAATGAGACCACTGAGGGGGAGCACCGGGGTCTGAGACCTC 413
QY 393 CGAGTACCCCGGACCTCCATACCCCATTTCC 425
|||||
Db 414 CCAGTACCCCGGACCTCCATACCCCATTTCC 446

RESULT 12
BO930450 954 bp mRNA linear EST 20-AUG-2002
LOCUS BO930450 8946139 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6462494
DEFINITION 5', mRNA sequence.
ACCESSION BO930450
VERSION BO930450.1 GI:22345481
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 954)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LINC2652 row: h column: 15
High quality sequence stop: 576.
Location/Qualifiers
1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6462494"
/clone_lib="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned

BASE COUNT	205 a	320 c	470 g
ORIGIN			

Query Match	89.5%;	Score 20071	
Best Local Similarity	99.7%;	Pred. No. 1.3e-83;	
Matches 381; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0

[illegible]

RESULT	13
BG576875	
LOCUS	BG576875
DEFINITION	60259923BF1 NIH_MCC_87 Homo sapiens CDNA clone IMAGE:4707753 5', mRNA sequence.

ACCESSION	BC576875	GI:13584528
VERSION	BC576875.1	
KEYWORDS	EST	
SOURCE	human	
ORGANISM	Homo sapiens	
	Chordata: Vertebrata: Euteleostomi	

REFERENCE
NIH-MGC <http://mgc.nci.nih.gov/>.
1 (bases 1 to 700)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Nathan Wolfe, et al.
TITLE Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
e-mail: strausbe-r@mai.nih.gov

Email: cgarcia@lml.uva.es
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov
plate: LLAM10578 row: b column: 10
High quality sequence stop: 673.

```
FEATURES
  source
    Location/Qualifiers
      1..700
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
```

```

/clone="IMAGE:4707753"
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally: oligo-dt primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

Query Match	87.1%;	Score 3/0;	DB 12;	Length 100;
Best Local Similarity	99.58;	Pred. NO. 4.4e-81;		
Matches 392;	Conservative	0;	Mismatches 0;	Indels 2; Gaps 2

QY	33	GGATGCTGCTCMAAGAGGGGCTCCGAGGGAGAGAGACCACCAATCTGTAGCTGGAACTGG	136
Db	84	GGATGGCTGCTCAAGAGGCGTCCGAGGGAGAGAGGCCCAATCTGTAGCTGGAAAGTTGG	143
QY	93	GAGCTCTCGAAGGGCCATTCATCTTCTCCGTGACTTGGTACAGCCAGCTGCTACCAATTTCT	152
Db	144	GAGCTTCGAGAGGCCATTCATCTTCTCCGTGACTTGGTACAGCCAGCTGCTACCAATTTCT	203
QY	153	GGAGATGTTTTCOCAGTATCCACAGGGCTCCCGCAGCCCGCAGTTTGGAGGAGGGGTG	212
Db	204	GGAGATGTTTTCOCAGTATCCACAGGGCTCCCGCAGCCCGCAGTTTGGAGGAGGGGTG	263
QY	213	CCGCAACAGCAGAGTCTCTGGACCTGACACAGGAGCCGCACTTTGGAACCCGGGGAGAG	272
Db	264	CCGCAACAGCAGAGTCTCTGGACCTGACACAGGAGCCGCACTTTGGAACCCGGGGAGAG	323
QY	273	AGCCAGGTGGCCACGATACCAATGGCAATTCATATGCAACGGCGGCTTTATGACTATCACT	332
Db	324	AGCCAGGTGGCCACGATACCAATGGCAATTCATATGCAACGGCGGCTTTATGACTATCACT	383
QY	333	GGCAAC-ATCTAATATACATGACACAGTACTGGGGGGAGCACCGGGTCTTGGAAAGCT	391
Db	384	GGCAACATCTAATATACATGACACAGTACTGGGGGGAGCACCGGGTCTTGGAAAGCT	443
QY	392	CCGAGTACCCCGGAACCTCATACCCATTTCC	425
Db	444	CCGAGTACCCCGGAACCTCATACCCATTTCC	476

RESULT 14	749 bp	mRNA	linear	EST 20-OCT-2000
BE789430				
LOCUS	BE789430			
DEFINITION	6014822671 NIH_MGC_68 Homo sapiens cDNA			
	mRNA sequence.			

ACCESSION	BE/89430	GI:10210628
VERSION	BE789430.1	
KEYWORDS	EST,	
SOURCE	human.	

ORGANISM	REFERENCE
<i>Homo sapiens</i>	1 (bases 1 to 749)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>,
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
email: cranbs@ncl.nih.gov

Email: cga@lml.
 Tissue Procurement: DCTD/DRP/Gazdar
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov
plate: LLM9659 row: e column: 17
High quality sequence stop: 743.

FEATURES

source 1. 749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3884920"
/issue="NIH-MGC-68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies." 167 a 241 c 220 g 121 t

BASE COUNT
ORIGIN

Query Match 83.5%; Score 355; DB 12; Length 749;
Best Local Similarity 99.7%; Pred. No. 2.2e-77;
Matches 366; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 60 GGAGAGGAGCCCAA-TCCCTAGCTGGAAGCTGGAGAGCTCCGAAGGCCATCCATCTT 118
Db 1 GGAGAGGAGCCCAAAGTCTGAGCTGGAAGCTGGAGAGCTCCGAAGGCCATCCATCTT 60

QY 119 CCCGACTGTGTACAGCCACTGCTACCATTTCTGAGATGTTTCCCATGATCCACTGG 178
Db 61 CCCGACTGTGTACAGCCACTGCTACCATTTCTGAGATGTTTCCCATGATCCACTGG 120

QY 179 GCTCCCGCAGCCCGCAGTTTGGAGGAGGAGGCTGCCCAAGAGAGTCTCTGGAGCT 238
Db 121 GCTCCCGCAGCCCGCAGTTTGGAGGAGGAGGCTGCCCAAGAGAGTCTCTGGAGCT 180

QY 239 GACCAGGAGCCGAGTGTGGAACCCGGGAGACAGAGCCAGGTGGCCACAGGTACCAATGG 298
Db 181 GACCAGGAGCCGAGTGTGGAACCCGGGAGACAGAGCCAGGTGGCCACAGGTACCAATGG 240

QY 299 CATTCATGTACCGCGGGGTGTATGACTATCATCTGCGCAACATCTACATCAATGAGAC 358
Db 241 CATTCATGTACCGCGGGGTGTATGACTATCATCTGCGCAACATCTACATCAATGAGAC 300

QY 359 AGTACTGGGGGAGCACCGGGGTCTGGAGACTCCGAGCTCCCGGCAAGCTCCATGCC 418
Db 301 AGTACTGGGGGAGCACCGGGGTCTGGAGACTCCGAGCTCCCGGCAAGCTCCATGCC 360

QY 419 CATTCGCC 425
Db 361 CATTCGCC 367

RESULT 15
BE207591 579 bp mRNA linear EST 27-JUN-2000
LOCUS BE207591
DEFINITION bb66f10.y1 NIH-MGC.9 Homo sapiens cDNA clone IMAGE:3030475 5'
similar to gb:U04270 TUMOR NECROSIS FACTOR RECEPTOR 2 RELATED
PROTEIN PRECURSOR (HUMAN); gb:U38423 Mus musculus lymphotoxin-beta
receptor gene, complete (MOUSE);, mRNA sequence.
ACCESSION BE207591
VERSION BE207591.1 GI:8750989
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 579)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apds@email.nih.gov
Tissue Procurement: DCTD/DFP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

image.lnl.gov/image/html/resources.shtml
Seq primer: -40RP from Gibco
High quality sequence stop: 427.
Location/Qualifiers
source 1. 579
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3030475"
/issue="NIH-MGC-9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pORF; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 100 a 198 c 152 g 127 t 2 others
ORIGIN

Query Match 81.6%; Score 346.6; DB 10; Length 579;
Best Local Similarity 92.1%; Pred. No. 2.4e-75;
Matches 386; Conservative 0; Mismatches 31; Indels 2; Gaps 2;

QY 8 AGAGCCACTGCCCGCAGAGATGTCAGATGCTCTCAAGAGGCTCCGAGGAGAGGG 67
Db 20 AGAGCCACTGCTCTCTGCGGAATGAGATGCTCTCAAGA-GGCTCCGAGGAGAGGG 78

QY 68 ACCCAATCTGTACCTGGAAGCTGGAGCCCTCCGAAGGCCCATCACTTCCCTGACTT 127
Db 79 ACCCAATCTGTGTCTGGAAGCTGGAGCCCTCCGAAGGCCCATCACTTTCCTGACTT 138

QY 128 GGTACAGCAGCTGCTACCCAT-TTCTGAGATGTTTCCCGATATCACTGGGCTCCCG 186
Db 139 GGTACAGCAGCTGTTGCCCGCTGTGGAGATGTTTCCCGATATCACTGGGCTCCCG 198

QY 187 CAGCCCAAGTTTGGAGGAGGAGGCTCCGAAGAGATGCTGAGCTGACCAAGGG 246
Db 199 CAGCCCAAGTTTGGAGGAGGAGGCTCCGAAGAGATGCTGAGCTGACCAAGGG 258

QY 247 AGCCGAGTTGGAACCCGGGAGAGAGCCAGTGGCCACAGGTACCAATGGCAATTCATG 306
Db 259 AGCCGAGTTGGAACCCGGGAGAGAGCCAGTGGCCACAGGTACCAATGGCAATTCATG 318

QY 307 TCACCGGGGGTCTATGACTATCACTGAGCAACATCTACATCAATGAGACCACTACTGG 366
Db 319 TCACCGGGGGTCTATGACTATCACTGAGCAACATCTCTACATGAGACCAATTCCTGG 378

QY 367 GGGAGCACCGGGTCTGAGAGACCTCCAGTACCCCGAAGCCCATACCCCAATTCGCC 425
Db 379 GGGAGCACCGGGTCTGAGAGACCTCCAGTACCCCGAAGCCCATACCCCAATTCGCC 437

Search completed: April 15, 2003, 16:26:35
Job time : 495.898 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:11:20 ; Search time 13.6946 Seconds
(without alignments)
9517.464 Million cell updates/sec

Title: US-09-917-372-2_COPY_853_1277

Perfect score: 425

Sequence: 1 atccatgagcaccatgcc.....aacctcatcaccatccc 425

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PTDS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34.2	8.0	268	1	US-08-039-137-16 Sequence 16, Appl
2	33	7.8	1057	3	US-09-188-930-18 Sequence 18, Appl
3	33	7.8	1722	4	US-09-142-623-14 Sequence 14, Appl
4	33	7.8	2004	3	US-09-188-930-230 Sequence 230, App
5	32.8	7.7	376	4	US-09-149-476-201 Sequence 201, App
6	32.2	7.6	2861	4	US-08-482-073-10 Sequence 10, Appl
7	32.2	7.6	3647	1	US-07-914-281-7 Sequence 7, Appl
8	32.2	7.6	3647	1	US-08-393-246-7 Sequence 7, Appl
9	32.2	7.6	3647	1	US-08-525-058A-7 Sequence 7, Appl
10	32.2	7.6	3647	2	US-08-696-731-7 Sequence 7, Appl
11	32.2	7.6	3647	4	US-09-042-531-7 Sequence 7, Appl
12	32.2	7.6	3647	5	PCT-US91-00899-4 Sequence 4, Appl
13	32.2	7.6	4403765	4	US-09-103-840A-2 Sequence 5, Appl
14	31.8	7.5	1977	4	US-09-548-372D-5 Sequence 5, Appl
15	31.8	7.5	1977	4	US-09-548-372D-5 Sequence 5, Appl
16	31.8	7.5	2070	4	US-09-548-372D-3 Sequence 3, Appl
17	31.8	7.5	2070	4	US-09-548-372D-3 Sequence 3, Appl
18	31.8	7.5	36741	4	US-09-548-367D-3 Sequence 3, Appl
19	31.6	7.4	550	4	US-09-301-665-3 Sequence 3, Appl
20	31.6	7.4	550	4	US-09-149-476-19 Sequence 19, Appl
21	31	7.3	269	1	US-08-050-073-2 Sequence 2, Appl
22	31	7.3	269	1	US-08-050-073-11 Sequence 11, Appl
23	31	7.3	269	1	US-08-050-073-14 Sequence 14, Appl
24	31	7.3	269	1	US-08-050-073-38 Sequence 38, Appl
25	31	7.3	1244	3	US-08-463-903-23 Sequence 23, Appl
26	31	7.3	1244	4	US-07-935-695-23 Sequence 23, Appl
27	31	7.3	4403765	4	US-09-103-840A-2 Sequence 2, Appl

28	30.4	7.2	1930	4	US-09-724-864-9 Sequence 9, Appl
29	30.4	7.2	1953	3	US-08-826-246-1 Sequence 1, Appl
30	30.4	7.2	1953	3	US-08-944-495-1 Sequence 1, Appl
31	30.4	7.2	1953	3	US-09-126-640-1 Sequence 1, Appl
32	30.4	7.2	1953	4	US-08-925-588-1 Sequence 1, Appl
33	30.4	7.2	1953	4	US-09-288-292A-1 Sequence 1, Appl
34	30.4	7.2	4086	7	US-08-313-181-1 Sequence 1, Appl
35	30.4	7.2	43950	4	US-09-735-934A-3 Sequence 3, Appl
36	30.2	7.1	2070	1	US-08-394-326-1 Sequence 1, Appl
37	30.2	7.1	2070	3	US-09-082-306-1 Sequence 1, Appl
38	30	7.1	1710	1	US-07-903-103-3 Sequence 3, Appl
39	30	7.1	1710	1	US-08-044-619A-3 Sequence 3, Appl
40	30	7.1	1710	1	US-08-283-911-3 Sequence 3, Appl
41	30	7.1	1710	1	US-08-245-500A-4 Sequence 4, Appl
42	30	7.1	1710	1	US-08-390-546-4 Sequence 4, Appl
43	30	7.1	1710	1	US-08-390-479A-4 Sequence 4, Appl
44	30	7.1	1710	1	US-08-357-393-4 Sequence 4, Appl
45	30	7.1	1710	1	US-08-390-516C-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-039-137-16
Sequence 16, Application US/08039137
Patent No. 5759771
GENERAL INFORMATION:
APPLICANT: Tilius J.G., Marcel
TITLE OF INVENTION: Method of Determining a Genotype by
TITLE OF INVENTION: Comparing the Nucleotide Sequence of Members of a Gene
Patent No. 5759771
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSER: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/039,137
FILING DATE: 14-Apr-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 0550-0024.10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: DRI cDNA, Fig. 1
FEATURE:
NAME/KEY: CDS
LOCATION: 1..267
US-08-039-137-16

```
Query Match      8.0%; Score 34.2; DB 1; Length 268;
Best Local Similarity 54.3%; Pred. No. 0.5;
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 207 GGGGCGCCGACAGACAGATCCTGTGACCTGACAGGAGCCGCGAGTTGGAAACCGGG 266
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 GCGGTGACGAGAGTGGGGCGGGCTGATGCCGAGTACTGGAACGCGACAGAACCTCCTG 189

QY 267 GAGCAGAGCCAGGTGGCCCGAGGTACCAATGCAATGTCACCGGGGGTCTATGACT 326
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 GAGCAGAGCGCGCGCGGTGACACCTACTGACAGACACAACTACGGGGTGTGAGAGC 249

QY 327 ATCACTG 333
      |||||
Db 250 TTCACAG 256

RESULT 2
US-09-188-930-18
: Sequence 18, Application US/09188930A
: Patent No. 6150502
: GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Murison, Rene
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: TITLE OF INVENTION: and Methods For Their Use
: FILE REFERENCE: 11000.1011c1
: CURRENT APPLICATION NUMBER: US/09/188,930A
: CURRENT FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: PastSeq for Windows Version 3.0
: SEQ ID NO 18
: LENGTH: 1057
: TYPE: DNA
: ORGANISM: Rat
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (45)...(45)
: NAME/KEY: unsure
: LOCATION: (53)...(53)
: NAME/KEY: unsure
: LOCATION: (116)...(116)
: NAME/KEY: unsure
: LOCATION: (118)...(118)
: US-09-188-930-18

Query Match      7.8%; Score 33; DB 3; Length 1057;
Best Local Similarity 51.0%; Pred. No. 1.7;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 53 TCCGAGGGAAGAGACCAATCTGTAGCTGGAAAGCTGGAGCCCGGAAGCCCATCC 112
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 353 TCCGAGCGCGTTGGATCCGATCGCAGCGGAGGAGTCCCATTAACGAGCTTCC 412

QY 113 ATACTCCCTGACTTGTAGACGACCACTGTACCATTTCTGGAATGTTTCCCAATTC 172
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 413 CTAAAGCCCTGGAGAGTTGATCTCCAAAGTACTCTCTCGGGAGCTACACCTGTATCA 472

QY 173 CACTGGGCTCCCGAGCCGAGTTTGGAGGC 205
      ||||| ||||| ||||| ||||| |||||
Db 473 CGCAAGGCTTTGGAGGACCGGATTACTGGGGC 505

RESULT 3
US-09-142-623-14/c
: Sequence 14, Application US/09142623
: Patent No. 6337201
: GENERAL INFORMATION:
: APPLICANT: Koji YANAI et al.
: TITLE OF INVENTION: -FRUCTOFURANOSIDASE AND ITS GENE, METHOD OF

TITLE OF INVENTION: ISOLATING -FRUCTOFURANOSIDASE GENE, SYSTEM FOR PRODUCING
TITLE OF INVENTION: -FRUCTOFURANOSIDASE, AND -FRUCTOFURANOSIDASE VARIANT
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESS: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/142,623
FILING DATE: September 10, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 98-0989*/LC(WMC)/144
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1722 base pairs
TYPE: Nucleic acid
STRANDEDNESS: Double stranded
TOPOLOGY: Linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Microorganism: Scopulariopsis brevicaulis IF04843
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1...1722
IDENTIFICATION METHOD: E
US-09-142-623-14

Query Match      7.8%; Score 33; DB 4; Length 1722;
Best Local Similarity 53.5%; Pred. No. 2;
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 296 TGGATTGATGTACACCGCGGGTCTATGACTATGCACTGGCAACATCTACATCAATGG 355
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 838 TGGCAAAAGATTTCACCGTGGGGTGTGAGCCATATCTGCAAGACCGCAATGTTGATGA 779

QY 356 ACCAGTACTGGGGGAGACACCGGGTCTGAGAGACTCCAGACTACCCCGAAGCTGCATA 415
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 778 CCTGAAATTTATACCCCGACCGCCAGCCAGTCACCGCTGCCCGAGGTGAGTTCCCTT 719

QY 416 CCCCATTC 424
      |||||
Db 718 CCTGCTTC 710

RESULT 4
US-09-188-930-230
: Sequence 230, Application US/09188930A
: Patent No. 6150502
: GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Murison, Rene
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: TITLE OF INVENTION: and Methods For Their Use
: FILE REFERENCE: 11000.1011c1
```


Wed Apr 16 07:59:41 2003

us-09-9i7-372-2_copy_853_1277.rni

```

; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 230
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Rat
US-09-188-930-230

Query Match      7.8%; Score 33; DB 3; Length 2004;
Best Local Similarity 51.0%; Pred. No. 2.1;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY  53  TCCGCAGGAGGAGGACCAATCCTTACTGTGAGAGCCCTCCGAGGCCATCC 112
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
DB  287  TCCGCAGCGTTGGGATTCGATCTGCAGCGAGAGAGTGTCCATTACAGGCTTTCC 346
      || |||| | || || || || || || || || || || || || || || ||
QY  113  ATACTTCCTGACTGTTGACAGCACTGCTACCCATTTCGTGAGATGTTCCCACTATC 172
      || |||| | || || || || || || || || || || || || || || ||
DB  347  CTAAAGCCCTGGAGACAGTGTGATCTCCAGTACTCTGTGGGAGGACTACATTCAC 406
      || |||| | || || || || || || || || || || || || || || ||
QY  173  CACTGGGCTCCCGCAGCCCACTTTGAGGC 205
      | | |||| | || || | || || || || || || || || || || || ||
DB  407  CGCAAGCTTTTGAGAGCCCGATCTACTGGGGC 439
      | | |||| | || || | || || || || || || || || || || || ||

RESULT 5
US-09-149-476-201/c
; Sequence 201, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
```

EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 879
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 880
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 894
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 911
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 636
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 874
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 910
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 864
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 631
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 845
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 892
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/057, 761
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/047, 595
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047, 599
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047, 588
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047, 585
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047, 586
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047, 590
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047, 594
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047, 589
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047, 593
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/047, 614
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/043, 578
 EARLIER FILING DATE: 1997-04-11
 EARLIER APPLICATION NUMBER: 60/043, 576
 EARLIER FILING DATE: 1997-04-11
 EARLIER APPLICATION NUMBER: 60/047, 501
 EARLIER FILING DATE: 1997-05-23
 EARLIER APPLICATION NUMBER: 60/043, 670
 EARLIER FILING DATE: 1997-04-11
 EARLIER APPLICATION NUMBER: 60/056, 632
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 664
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 876
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 881
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 909
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 875
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 862
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 887
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/056, 908
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/048, 964
 EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/057, 650
 EARLIER FILING DATE: 1997-09-05
 EARLIER APPLICATION NUMBER: 60/056, 884
 EARLIER FILING DATE: 1997-08-22
 EARLIER APPLICATION NUMBER: 60/057, 669
 EARLIER FILING DATE: 1997-09-05
 EARLIER APPLICATION NUMBER: 60/049, 610
 EARLIER FILING DATE: 1997-06-13
 EARLIER APPLICATION NUMBER: 60/061, 060
 EARLIER FILING DATE: 1997-10-02

Query Match 7.7%; Score 32.8; DB 4; Length 376;
 Best Local Similarity 47.5%; Pred. No. 1.4;
 Matches 94; Conservative 1; Mismatches 103; Indels 0; Gaps 0;

QY 62 AGAGGACCAATCTCTAGCTGAGAGGAGCCGACGAGCCATTCATCTTCC 121
 DB 354 AATGAGCCAGATGCTTATGAGCTGCTCCAGAGCCATTCGATCAGATCCCTT 295
 QY 122 TGACTGTACAGCCACTGCTACCAATTTCTGAGATGTTCCCATTCACCTGGGCT 181
 DB 294 AATGAGCCAGATGCTTATGAGCTGCTCCAGAGCCATTCGATCAGATCCCTT 235
 QY 182 CCCCAGCCCACTTTTGGAGGAGGAGGCGCCGACAGAGAGCTCTGAGCTGAC 241
 DB 234 CACAGGAGCCCTGGCTGATGGGCTCTGCTGGGAGTACAGATGCTCAGTCCAGATC 175
 QY 242 CAGGAGCCGAGTTGA 259
 DB 174 TTCAGAGACCTCTAGGA 157

RESULT 6
 US-08-482-073-10/c
 Sequence 10, Application US/08482073
 Patent No 6307025
 GENERAL INFORMATION:
 APPLICANT: Hession, Catherine A.
 APPLICANT: Lobb, Roy R.
 APPLICANT: Goelz, Susan E.
 APPLICANT: Osborn, Laurelee
 APPLICANT: Benjamin, Christopher D.
 APPLICANT: Rosa, Margaret D.
 TITLE OF INVENTION: ENDOTHELIAL CELL-LEUKOCYTE ADHESION
 TITLE OF INVENTION: MOLECULES (ELAMS) AND MOLECULES INVOLVED IN LEUKOCYTE
 TITLE OF INVENTION: ADHESION (MILAS)
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Neave
 STREET: 1251 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States of America
 ZIP: 10020
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/482,073
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/486,336
 FILING DATE:
 APPLICATION NUMBER: US 07/608298
 FILING DATE: 31-OCT-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/US 90/02357
 FILING DATE: 27-APR-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/452675

FILING DATE: 18-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/359516
FILING DATE: 01-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/354151
FILING DATE: 28-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B124C1P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2861 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-482-073-10

Query Match 7.6%; Score 32.2; DB 4; Length 2861;
Best Local Similarity 57.4%; Pred. No. 4;
Matches 58; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 179 GCTCCCCGAGCCCGAGTTTGGAGGCGAGGTCGCCAGACAGACAGATCTCTGGACCT 238
DB 447 GCTCGCCGCTGGCCCGGAGTGTGAAGGGGCGGTCCCAAGATGCAAGGCGCGGCCCT 388
QY 239 GACCAAGGAGCGCGAGTTGGAAACCCGAGGAGCAGACAGCCAGG 279
DB 387 GCCCTGCTCTCCCAAGTGCAGGCGGCGGAGCCGCCAGG 347

RESULT 7

US-07-914-281-7/C
Sequence 7, Application US/07914281
Patent No. 5324663

GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/914,281
FILING DATE: 19920720
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Lavelle, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 3647 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
US-07-914-281-7

Query Match 7.6%; Score 32.2; DB 1; Length 3647;
Best Local Similarity 57.4%; Pred. No. 4.3;
Matches 58; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 179 GCTCCCCGAGCCCGAGTTTGGAGGCGAGGTCGCCAGACAGACAGATCTCTGGACCT 238
DB 1840 GCTCGCCGCTGGCCCGGAGTGTGAAGGGGCGGTCCCAAGATGCAAGGCGCGGCCCT 1781
QY 239 GACCAAGGAGCGCGAGTTGGAAACCCGAGGAGCAGACAGCCAGG 279
DB 1780 GCCCTGCTCTCCCAAGTGCAGGCGGCGGAGCCGCCAGG 1740

RESULT 8

US-08-393-246-7/C
Sequence 7, Application US/08393246
Patent No. 5595900

GENERAL INFORMATION:
APPLICANT: LOWE, JOHN B.
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Fourth Floor
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,246
FILING DATE:
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,433
FILING DATE: 30-MAR-1994
APPLICATION NUMBER: US 07/914,281
FILING DATE: 20-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lavelle, Jean-Paul M. P.
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-060-55
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3647 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
US-08-393-246-7

RESULT 10
US-08-696-731-7/c

1 Patent No. 6268193
 2
 3 GENERAL INFORMATION:
 4
 5 APPLICANT: LOWE, JOHN B.
 6
 7 TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
 8 OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
 9
 10 TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
 11
 12 TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
 13
 14
 15

```

ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/00899
FILING DATE: 19910214
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lavalleye Ph.D., Jean-Paul
REGISTRATION NUMBER: 31,451
REFERENCE/DOCKET NUMBER: 2363-021-55 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-5940
TELEFAX: (703)486-2347
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3647 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
PCT-US91-00899-4

Query Match          7.6%: Score 32.2; DB 5; Length 3647;
Best Local Similarity 57.4%; Pred. No. 4.3; Mismatches 43; Indels 0; Gaps 0;
Matches 58; Conservative 0;

QY 179 GCTCCCCGACAGCCCGACTTTTGGAGCAGAGGGGTGCCCAACAGCAGAGTCCTCTGAGCCT 238
      ||| |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Db 1840 GCTCCGCCTGCGCCGCGGAATGGAAGGGGGCGGTCCCAAGAATCAAAGCCGCGGCCCT 1781

QY 239 GACCAGGAGCGCCGCGAGTGGAACCCGGGGAGCAGAGCCAGG 279
      ||| | || | || | || | |||| | |||| | |||| |
Db 1780 GCCCTGCTCTCCCAAGTCCGCGGGGCGGAGCCGCGCAGG 1740

RESULT 13
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VEMTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match          7.6%: Score 32.2; DB 4; Length 4403765;
Best Local Similarity 48.9%; Pred. No. 31;
Matches 114; Conservative 0; Mismatches 118; Indels 1; Gaps 1;

```

Search completed: April 15, 2003, 16:46:55
Job time : 1040.69 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 15:24:11 ; Search time 29.6131 Seconds
(without alignments)
12588.918 Million cell updates/sec

Title: US-09-917-372-2_COPY_853_1277
Perfect score: 425
Sequence: 1 atccattagagccacgtccccc.....aacctccataccattccc 425

Scoring table: IDENTITY NUC
Gapop:10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2-6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2-6/ptodata/2/pubpna/PCIT_NEM_PUB.seq:*
3: /cgn2-6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4: /cgn2-6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
5: /cgn2-6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
6: /cgn2-6/ptodata/2/pubpna/PCITUS_PUBCOMB.seq:*
7: /cgn2-6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
8: /cgn2-6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2-6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
10: /cgn2-6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2-6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
12: /cgn2-6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2-6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
14: /cgn2-6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	425	100.0	425	10	US-09-907-372-9
2	425	100.0	1982	10	US-09-907-372-2
3	423.4	99.6	574	10	US-09-907-372-8
4	403	94.8	651	10	US-09-907-372-7
5	393	92.5	2136	9	US-09-934-531-1348
6	393	92.5	2136	10	US-09-962-436-262
7	393	92.5	2136	10	US-09-880-107-2135
8	253.8	59.7	281	10	US-09-867-701-5601
9	251.2	59.1	289	10	US-09-867-701-5666
10	171.8	40.4	371	10	US-09-907-372-16
11	135.8	32.0	206	10	US-09-907-372-13
12	98.4	23.2	219	10	US-09-907-372-11
13	71	16.7	279	10	US-09-907-372-11
14	47.8	11.2	862	10	US-09-907-372-12
15	35.4	8.3	809	10	US-09-844-864-22
16	34	8.0	6788	9	US-09-870-759-101
17	33.6	7.9	33675	10	US-09-921-992-2
18	33	7.8	1057	9	US-10-152-661-18
19	33	7.8	1057	9	US-09-866-050A-18

C	20	33	7.8	1722	9	US-09-990-385-14	Sequence 14, Appl
	21	33	7.8	1890	9	US-10-152-661-447	Sequence 447, App
	22	33	7.8	1890	9	US-09-866-050A-447	Sequence 447, App
	23	33	7.8	2004	9	US-10-152-661-230	Sequence 230, App
	24	33	7.8	2004	9	US-09-866-050A-230	Sequence 230, App
C	25	32.8	7.7	376	9	US-09-809-391-201	Sequence 201, App
	26	32.2	7.6	785	9	US-09-902-941-1868	Sequence 1868, Ap
	27	32.2	7.6	785	9	US-09-849-626-1868	Sequence 1868, Ap
	28	32.2	7.6	785	9	US-10-017-754-1868	Sequence 1868, Ap
C	29	32.2	7.6	3647	10	US-09-863-475A-7	Sequence 7, Appl
	30	31.8	7.5	671	9	US-10-184-644-346	Sequence 346, App
	31	31.8	7.5	1911	9	US-09-470-954A-44	Sequence 44, Appl
	32	31.8	7.5	1977	10	US-09-794-927-5	Sequence 5, Appl
	33	31.8	7.5	1977	10	US-09-794-927-5	Sequence 5, Appl
	34	31.8	7.5	1977	10	US-09-794-743-5	Sequence 5, Appl
	35	31.8	7.5	1977	10	US-09-794-748-5	Sequence 5, Appl
	36	31.8	7.5	1977	10	US-09-794-925-5	Sequence 5, Appl
	37	31.8	7.5	1977	10	US-09-681-442-5	Sequence 5, Appl
	38	31.8	7.5	2070	10	US-09-794-927-3	Sequence 3, Appl
	39	31.8	7.5	2070	10	US-09-795-847-3	Sequence 3, Appl
	40	31.8	7.5	2070	10	US-09-794-743-3	Sequence 3, Appl
	41	31.8	7.5	2070	10	US-09-794-748-3	Sequence 3, Appl
	42	31.8	7.5	2070	10	US-09-794-925-3	Sequence 3, Appl
	43	31.8	7.5	2070	10	US-09-681-442-3	Sequence 3, Appl
	44	31.8	7.5	3252	9	US-09-795-903A-1	Sequence 1, Appl
	45	31.8	7.5	3252	10	US-09-796-264-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-907-372-9/c
Sequence 9, Application US/09907372
Patent No. US20020068242A1
GENERAL INFORMATION:
APPLICANT: Lal, Preethi G.
TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
FILE REFERENCE: PC-0050 US
CURRENT APPLICATION NUMBER: US/09/907.372
CURRENT FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 425
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID NO. US20020068242A1 7716340H1
US-09-907-372-9

Query Match	100.0%	Score 425:	DB 10:	Length 425:
Best Local Similarity	100.0%	Pred. No. 6.5e-124:		
Matches 425:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
OY	1	ATCCATTAGAGCCACTGCCCCCAGAGATGTCAGATGCTGCTCAAGAGCGCTCCGACG	60	
DB	425	ATCCATTAGAGCCACTGCCCCCAGAGATGTCAGATGCTGCTCAAGAGCGCTCCGACG	366	
OY	61	GAGAGGAGCCCAATCTGAGCTGGAAGCTGGAGCTCCGAAGGCCATCATCTTCC	120	
DB	365	GAGAGGAGCCCAATCTGAGCTGGAAGCTGGAGCTCCGAAGGCCATCATCTTCC	306	
OY	121	CTGACTTGTTAGAGCCACTGCTATCCATTCTTGAGATGTTTCCCATGATATCCATGAGC	180	
DB	305	CTGACTTGTTAGAGCCACTGCTATCCATTCTTGAGATGTTTCCCATGATATCCATGAGC	246	
OY	181	TCGCCGACAGCCCACTTTTGGAGGAGGGGTCGCCACAGAGAGATCTTGACCTGA	240	
DB	245	TCGCCGACAGCCCACTTTTGGAGGAGGGGTCGCCACAGAGATCTTGACCTGA	186	

```

QY 241 CCAGGGAGCCGCGAGTTGGAACCCGGGGAGACAGCCAGGTTGGGCCACGGTACCAATGSGA 300
    |||||
Db 185 CCAGGGAGCCGCGAGTTGGAACCCGGGGAGACAGCCAGGTTGGGCCACGGTACCAATGSGA 126
QY 301 TTCATGTACACCGGGGGTGTATGACTATGCTAGCTGGCAACATCTACATCTAACATGAGCCAG 360
    |||||
Db 125 TTCATGTACACCGGGGGTGTATGACTATGCTAGCTGGCAACATCTACATCTAACATGAGCCAG 66
QY 361 TACTGGGGGAGACACCGGGTCTGGAGACCTCCGAGCTCCCGGACCTCCGAGCTCCATACCCCA 420
    |||||
Db 65 TACTGGGGGAGACACCGGGTCTGGAGACCTCCGAGCTCCCGGACCTCCGAGCTCCATACCCCA 6
QY 421 TTCCC 425
    |||||
Db 5 TTCCC 1

RESULT 2
US-09-907-372-2
; Sequence 2, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1982
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7497867CBI
US-09-907-372-2

Query Match          100.0%; Score 425; DB 10; Length 1982;
Best Local Similarity 100.0%; Pred. No. 9,1e-124;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCCATTAGAGCCACTGCCCCCAAGATGTAGAGATGCTGCTCTCAAGAGCGCTCCGAGG 60
    |||||
Db 853 ATCCATTAGAGCCACTGCCCCCAAGATGTAGAGATGCTGCTCTCAAGAGCGCTCCGAGG 912
QY 61 GAGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGCCTCCGAAGGCCCATCCATCTCC 120
    |||||
Db 913 GAGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGCCTCCGAAGGCCCATCCATCTCC 972
QY 121 CTGACTTGTTACAGCCACTGCTACCATTTCTGGAGATGTTTCCCAAGTATCCACTGGG 180
    |||||
Db 973 CTGACTTGTTACAGCCACTGCTACCATTTCTGGAGATGTTTCCCAAGTATCCACTGGG 1032
QY 181 TCCCCGAGGCCCGTGGTGGAGGCGGGGCTCCGCAACAGCAGATCTCTGGAGCTGA 240
    |||||
Db 1033 TCCCCGAGGCCCGTGGTGGAGGCGGGGCTCCGCAACAGCAGATCTCTGGAGCTGA 1092
QY 241 CCAGGGAGCCGCGAGTTGGAACCCGGGGAGACAGCCAGTGGGCCACGGTACCAATGSGA 300
    |||||
Db 1093 CCAGGGAGCCGCGAGTTGGAACCCGGGGAGACAGCCAGTGGGCCACGGTACCAATGSGA 1152
QY 301 TTCATGTACACCGGGGGTGTATGACTATGCTAGCTGGCAACATCTACATCTAACATGAGCCAG 360
    |||||
Db 1153 TTCATGTACACCGGGGGTGTATGACTATGCTAGCTGGCAACATCTACATCTAACATGAGCCAG 1212
QY 361 TACTGGGGGAGACACCGGGTCTGGAGACCTCCGAGCTCCCGGACCTCCGAGCTCCATACCCCA 420
    |||||
Db 1213 TACTGGGGGAGACACCGGGTCTGGAGACCTCCGAGCTCCCGGACCTCCGAGCTCCATACCCCA 1272
QY 421 TTCCC 425
    |||||
```

```

Db 1273 TTCCC 1277

RESULT 3
US-09-907-372-8
; Sequence 8, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 8234468H1
US-09-907-372-8

Query Match          99.8%; Score 423.4; DB 10; Length 574;
Best Local Similarity 99.8%; Pred. No. 2,2e-123;
Matches 424; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCATTAGAGCCACTGCCCCCAAGATGTAGAGATGCTGCTCTCAAGAGCGCTCCGAGG 60
    |||||
Db 50 ATCCATTAGAGCCACTGCCCCCAAGATGTAGAGATGCTGCTCTCAAGAGCGCTCCGAGG 109
QY 61 GAGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGCCTCCGAAGGCCCATCCATCTCC 120
    |||||
Db 110 GAGAGGAGCCCAATCTGTAGCTGGAAGCTGGAGCCTCCGAAGGCCCATCCATCTCC 169
QY 121 CTGACTTGTTACAGCCACTGCTACCATTTCTGGAGATGTTTCCCAAGTATCCACTGGG 180
    |||||
Db 170 CTGACTTGTTACAGCCACTGCTACCATTTCTGGAGATGTTTCCCAAGTATCCACTGGG 229
QY 181 TCCCCGAGGCCCGTGGTGGAGGCGGGTCCGCAACAGCAGATCTCTGGAGCTGA 240
    |||||
Db 230 TCCCCGAGGCCCGTGGTGGAGGCGGGTCCGCAACAGCAGATCTCTGGAGCTGA 289
QY 241 CCAGGGAGCCGCGAGTTGGAACCCGGGGAGACAGCCAGTGGGCCACGGTACCAATGSGA 300
    |||||
Db 290 CCAGGGAGCCGCGAGTTGGAACCCGGGGAGACAGCCAGTGGGCCACGGTACCAATGSGA 349
QY 301 TTCATGTACACCGGGGGTGTATGACTATGCTAGCTGGCAACATCTACATCTAACATGAGCCAG 360
    |||||
Db 350 TTCATGTACACCGGGGGTGTATGACTATGCTAGCTGGCAACATCTACATCTAACATGAGCCAG 409
QY 361 TACTGGGGGAGACACCGGGTCTGGAGACCTCCGAGCTCCCGGACCTCCGAGCTCCATACCCCA 420
    |||||
Db 410 TACTGGGGGAGACACCGGGTCTGGAGACCTCCGAGCTCCCGGACCTCCGAGCTCCATACCCCA 469
QY 421 TTCCC 425
    |||||
Db 470 TTCCC 474
```

```

RESULT 4
US-09-907-372-7/c
; Sequence 7, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
```



```

; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7716364H1
; us-09-907-372-7

```

```

Query Match          94.8%; Score 403; DB 10; Length 651;
Best Local Similarity 99.5%; Pred. No. 5,7e-117;
Matches 425; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```

```

QY 1 ATCCATTNAGCCACTGCCCCCAGAGATGTCAGAGATCGCTGCTCAAGAGCGCTCCGACAG 60
DB 427 ATCCATTNAGCCACTGCCCCCAGAGATGTCAGAGATCGCTGCTCAAGAGCGCTCCGACAG 368
QY 61 GAGAGGAGCCCAATCCTGAGCTGGAAGCTGGAGGCTCCGAAGGCCCATCCATCTTCC 120
DB 367 GAGAGGAGCCCAATCCTGAGCTGGAAGCTGGAGGCTCCGAAGGCCCATCCATCTTCC 308
QY 121 CTGACTTGTAGACCCACTGCTACCCATTTCTGAGAGATGTTCCCAAGATCCACTGAGC 180
DB 307 CTGACTTGTAGACCCACTGCTACCCATTTCTGAGAGATGTTCCCAAGATCCACTGAGC 248
QY 181 TCCCGCCAGCCCACTTTTGGAGGAGGAGGCTCCGACAGAGAGAGTCTCTGGACCTGA 240
DB 247 TCCCGCCAGCCCACTTTTGGAGGAGGAGGCTCCGACAGAGAGTCTCTGGACCTGA 188
QY 241 CCAGGAGGCGGAGTGTGAACCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
DB 187 CCAGGAGGCGGAGTGTGAACCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 128
QY 301 TTCATGTCAACCGGCGGTCTATGATCATCTGCAATCATCTACATCTCAATGAGACAG 360
DB 127 TTCATGTCAACCGGCGGTCTATGATCATCTGCAATCATCTACATCTCAATGAGACAG 68
QY 361 TACTGGGGGAGGACCGGGTCTCTGGAGACCTCC-AGCTACCCCGGAGACCTCC-ATACCC 418
DB 67 TACTGGGGGAGGACCGGGTCTCTGGAGACCTCC-AGCTACCCCGGAGACCTCC-ATACCC 8
QY 419 CATTCOC 425
DB 7 CATTCOC 1

```

RESULT 5

```

; Sequence 1348, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cano
; FILE REFERENCE: 689290-77
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1348
; LENGTH: 2136

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; us-09-954-531-1348

```

```

Query Match          92.5%; Score 393; DB 9; Length 2136;
Best Local Similarity 100.0%; Pred. No. 1e-113;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 33 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 92
DB 943 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1002
QY 93 GAGCCCTCCAGAGGCCCATCATCTTCCCTGACTTGTGACAGCACTGTACCATTTCT 152
DB 1003 GAGCCCTCCAGAGGCCCATCATCTTCCCTGACTTGTGACAGCACTGTACCATTTCT 1062
QY 153 GGAGATGTTCCCAATATCCACTGGGCTCCCGAGCCCACTTTTGGAGGAGGAGGAG 212
DB 1063 GGAGATGTTCCCAATATCCACTGGGCTCCCGAGCCCACTTTTGGAGGAGGAGGAG 1122
QY 213 CCGCAACAGCAGAGTCTCTGAGACTTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 272
DB 1123 CCGCAACAGCAGAGTCTCTGAGACTTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1182
QY 273 AGCCAGGTGGCCACGAGTACCAATGAGCATTTGATCAGGCGGAGTATGATCATCT 332
DB 1183 AGCCAGGTGGCCACGAGTACCAATGAGCATTTGATCAGGCGGAGTATGATCATCT 1242
QY 333 GGCACATCTACATCTACATGAGACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 392
DB 1243 GGCACATCTACATCTACATGAGACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1302
QY 393 CCAAGTACCCCGGAGACCTCATACCCCATTTCC 425
DB 1303 CCAAGTACCCCGGAGACCTCATACCCCATTTCC 1335

```

RESULT 6

```

; Sequence 262, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-75
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 262
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
; us-09-962-436-262

```

```

Query Match          92.5%; Score 393; DB 10; Length 2136;
Best Local Similarity 100.0%; Pred. No. 1e-113;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 33 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 92
DB 943 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1002
QY 93 GAGCCCTCCAGAGGCCCATCATCTTCCCTGACTTGTGACAGCACTGTACCATTTCT 152
DB 1003 GAGCCCTCCAGAGGCCCATCATCTTCCCTGACTTGTGACAGCACTGTACCATTTCT 1062
QY 153 GGAGATGTTCCCAATATCCACTGGGCTCCCGAGCCCACTTTTGGAGGAGGAGGAG 212

```

```

Db 1063 GGAGATGTTTCCCGATGATCCAGTGGGCTCCCGCAGAGCCCAATTTTGGAGGCGAGG 1122
OY 213 CCGCAAGACGAGATCTCTGGACCTGACGAGAGCCGAGTTGGAAACCCGGGAGGACAG 272
Db 1123 CCGCAAGACGAGATCTCTGGACCTGACGAGAGCCGAGTTGGAAACCCGGGAGGAG 1182
OY 273 AGCAGGTGGCCACGAGTACCATGATTCATGTCACCGGGGGTCTATGATATCACT 332
Db 1183 AGCAGGTGGCCACGAGTACCATGATTCATGTCACCGGGGGTCTATGATATCACT 1242
OY 333 GGCAACATCTACATCTACAAATGAGACAGTACTGGGGGAGACCGGGTCTTGAAGACCTC 392
Db 1243 GGCAACATCTACATCTACAAATGAGACAGTACTGGGGGAGACCGGGTCTTGAAGACCTC 1302
OY 393 CCAGCTACCCCGAAGCTCCATACCCATTTCCC 425
Db 1303 CCAGCTACCCCGAAGCTCCATACCCATTTCCC 1335

```

RESULT 7
US-09-880-107-2135

```

; Sequence 2135, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Scherf, Joseph G.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2135
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L04270
US-09-880-107-2135

```

Query Match 92.5%; Score 393; DB 10; Length 2136;
Best Local Similarity 100.0%; Pred. No. 1e-113;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 33 GGATCGCTGCTCAGAGCGCTCCGAGGAGAGGAGCCCAATCTCTAGCTGGAAGCTGG 92
Db 943 GGATCGCTGCTCAGAGCGCTCCGAGGAGAGGAGCCCAATCTCTAGCTGGAAGCTGG 1002
OY 93 GAGCCGCCAAGAGCGCCATCACTACTTCCCTGACTGTGACAGCACTGCTACCATTTCT 152
Db 1003 GAGCCGCCAAGAGCGCCATCACTACTTCCCTGACTGTGACAGCACTGCTACCATTTCT 1062
OY 153 GGAGATGTTTCCCGATGATCCAGTGGGCTCCCGCAGAGCCGAGTTTGGAGGAGGAG 212
Db 1063 GGAGATGTTTCCCGATGATCCAGTGGGCTCCCGCAGAGCCGAGTTTGGAGGAGGAG 1122
OY 213 CCGCAAGACGAGATCTCTGGACCTGACGAGAGCCGAGTTGGAAACCCGGGAGGAG 272
Db 1123 CCGCAAGACGAGATCTCTGGACCTGACGAGAGCCGAGTTGGAAACCCGGGAGGAG 1182
OY 273 AGCAGGTGGCCACGAGTACCATGATTCATGTCACCGGGGGTCTATGATATCACT 332
Db 1183 AGCAGGTGGCCACGAGTACCATGATTCATGTCACCGGGGGTCTATGATATCACT 1242
OY 393 GGCAACATCTACATCTACAAATGAGACAGTACTGGGGGAGACCGGGTCTTGAAGACCTC 392

```

```

Db 1243 GGCAACATCTACATCTACAAATGAGACAGTACTGGGGGAGACCGGGTCTTGAAGACCTC 1302
OY 393 CCAGCTACCCCGAAGCTCCATACCCATTTCCC 425
Db 1303 CCAGCTACCCCGAAGCTCCATACCCATTTCCC 1335

```

RESULT 8
US-09-867-701-5601

```

; Sequence 5601, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5601
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-5601

```

Query Match 59.7%; Score 253.8; DB 10; Length 281;
Best Local Similarity 97.1%; Pred. No. 3.1e-70;
Matches 269; Conservative 7; Mismatches 1; Indels 1;

```

OY 144 CCATTTCTGAGATGTTTCCCGATGATCCAGTGGGCTCCCGCAGAGCC-CAGTTTGA 202
Db 5 CCATTTCTGAGATGTTTCCCGATGATCCAGTGGGCTCCCGCAGAGCC-CAGTTTGA 64
OY 203 GCGAGGGGTGCGGACAGAGAGTCTTGACCTGACAGGAGAGCCGAGTTGGAAC 262
Db 65 GCGAGGGGTGCGGACAGAGAGTCTTGACCTGACAGGAGAGCCGAGTTGGAAC 124
OY 263 CCGGAGACGAGAGCGAGTGGCCCGACGATTCATGCTGACCGGGGCTAT 322
Db 125 CCGGAGACGAGAGCGAGTGGCCCGACGATTCATGCTGACCGGGGCTAT 184
OY 323 GACTATCACTGGCAACATCTACATCTACAAATGAGACAGTACTGGGGGAGACCGGGTCA 382
Db 185 GACTATCACTGGCAACATCTACATCTACAAATGAGACAGTACTGGGGGAGACCGGGTCA 244
OY 383 TGGAGACCTCCAGCTACCCCGAAGCTCCATACCC 419
Db 245 TGGAGACCTCCAGCTACCCCGAAGCTCCATACCC 281

```

RESULT 9
US-09-867-701-5666/c

```

; Sequence 5666, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5666
; LENGTH: 289
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-5666

```

Query Match	59.1%;	Score 251.2;	DB 10;	Length 289;
Best Local Similarity	98.5%;	Pred. NO. 2.1e-69;		
Matches 264; Conservative	0;	Mismatches 3;	Indels 1;	Gaps 1

QY 155 AGATGTTTCCCCAGTATCACTGGGCTCCCCGAGCCCGATTGAGCGAGGGGTGCC 214
|||||
Db 289 AGATGTTTCCCCAGTATCCACTGGGCTCCCCGAGCCCGATTGAGCGAGGGGTGCC 230

Dy 215 GCACACAGAGA-GTCTCTGGACCCTGACCAGGAGCCCGAGTTGGAAACCCGGGAGCAGA 273
|||||
Db 229 GCACACAGAGATGTCTCTTGACCTGACCAGGAGGCCCGGTTGGAAACCCGGGAGCAGA 170

Db 169 GCAGGTGGCCACCGGTACCAATGCATTCATGCACGGGGGTCTATGACTATCATCTG 110

Db 109 GCACATCTACATCTACATGGACCACTGCGGGGCGACCGGGTCTGGAGACCTCC 50

Db 49 CAGCTACCCCCGAACTTCATTTCCCCCT 22

US-09-907-372-16
; Sequence 16, Application US/09907372
; Patent No. US20020068242A1

```

; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT

```

```

;
; CURRENT APPLICATION NUMBER: 05/09/90/,312
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
;
SOFTWARE: PERL Program

```

```

;
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Canis familiaris

```

```

: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No. US20020068242A1 702245091H1
US-09-907-372-16

```

Query match	90.4%	Score	11.6	D5	10	Length	311
Best Local Similarity	81.0%	Pred. No.	1.86-4.1				
Matches 200; Conservative	0	Mismatches	47	Indels	0	Gaps	0

Db 4 GACTCCGACGACCCCACTCTGGAGAGAGAGATGCTACAGCAGAGATCTCTCAGCC 63

64 AGGCCAGAGAGCTGGAGCCTTGAGCTCCACAGAACAGGCCAGGCGGCCACACGGTACCAATG 123

Db	124	GCATTACACTCACCGCGGGTCTGTGACTGTGACTGGCAACATCTACATCTACAATGGGC	183
Oy	358	CAGTACTGGGGGGACACACCGGGTCTCTGGAGACCTCCCACTTACCCCGAAGCTCCATATCC	417

Dd	184	CAgTACTGGGGGGAGACACGAGGCCCTGGAGACCCCTGCTTCCCCAGAGCCTCCATACC	243
Qy	418	CCATTCC	424

Db 244 CCACCCC 250

RESULT 11
US-09-907-372-13
; Sequence 13, Application US/09907372
; Patent No. US20020068242A1

```

; APPLICANT: Lai, Pirelli G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US

```

```

; CURRENT FILING DATE: 2001-07-27
;
; NUMBER OF SEQ ID NOS.: 20
;
; SOFTWARE: PERL Program

```

```

; LENGTH: 206
; TYPE: DNA
; ORGANISM: Rattus norvegicus

```

NAME/REF: INUSC_124016
OTHER INFORMATION: Incyte ID No. US20020068242A1 700302531H11
US-09-907-372-13

Best Local	Similarly	Pred. No.	Matches	Conservative	Mismatches	Indels	Gaps
158	81.0%	3.1e-33	0	37	0	0	0

Db 2 CTGATCCAGGCCAGGAGCTGGAGGCTGGGGAACATGGCCAGGTGGCCACGGT 61

351 AATGACCAATCTACTGGGGGGGAGCCACCGGGTCTCTGGAGACTCCCACTACCCCGGAACTT 110
Db 62 GCGAATGGCATTACAGTGCAGCCGGAGGCTCTGTGACTGTCAACCGGCAATATCTACATATAC 122

Db 122 AATGGGCCAGTGCCTGGGGGGGAACACGGGGCCCTGGAGACCCTCCAGCTCCCCCTGAGCCT 181
QY 411 CCATACCCCAATTCCC 425

DB 182 CCA^TACC^CGACT^CCC 196

; Sequence 10, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:

```

1  APPLICANT:  Warren, Bridget A.
2  TITLE OF INVENTION:  TNF RECEPTOR 2 RELATED PROTEIN VARIANT
3  FILE REFERENCE:  PC-0050 US
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031

```

```

; SEQ ID NO 10
;
; SOFTWARE: PERL Program
;
; NUMBER OF SEQ ID NOS: 20
;
; CURRENT FILING DATE: 2001-07-27
;

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

OTHER INFORMATION: Incyte ID No. US20020068242A1 697459H1
NAME/KEY: unsure
LOCATION: 76, 131
OTHER INFORMATION: 3 + 3 at 34hr

OTHER INFORMATION: a, l, c, y, or other
US-09-907-372-10

Query Match	23.28;	Score 98.4;	DB 10;	Length 219;
Best Local Similarity	98.08;	Pred. No. 1.7e-21;		
Matches 99; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

325 CTATCACTGGCAACATCTACATCTACAATGGACCAGTACTGGGGGACCAACCGGTCTG 384

Db 9 CGATCTACTGGCAACATCTACATCTACATGACAGACAGTCTGGGGGAGACCAGCGGCTCTG 68
 QY 385 GAGACCTCCAGCTACACCCCGAAGCTCTCATATACCCCATTTCCC 425
 |||||||
 Db 69 GAGACCTNCCAGCTACCCCGAAGCTCTCATATACCCCATTTCCC 109

RESULT 13

US-09-907-372-11
 ; Sequence 11, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti G.
 ; APPLICANT: Warren, Bridget A.
 ; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
 ; FILE REFERENCE: PC-0050 US
 ; CURRENT APPLICATION NUMBER: US/09/907,372
 ; CURRENT FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 11
 ; LENGTH: 279
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: Incyte ID No. US20020068242A1 3321983H1
 ; US-09-907-372-11

Query Match

Best Local Similarity 16.7%; Score 71; DB 10; Length 279;
 Best Local Similarity 100.0%; Pred. No. 7,2e-13;
 Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 355 GACCACTACTGGGGGAGACACCGGCTCTGGAGACCTCCAGTACCCCGAAGCTCCAT 414
 |||||||
 Db 1 GACCACTACTGGGGGAGACACCGGCTCTGGAGACCTCCAGTACCCCGAAGCTCCAT 60
 QY 415 ACCCATTTCCC 425
 |||||||
 Db 61 ACCCATTTCCC 71

RESULT 14

US-09-907-372-12/c
 ; Sequence 12, Application US/09907372
 ; Patent No. US20020068242A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lal, Preeti G.
 ; APPLICANT: Warren, Bridget A.
 ; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
 ; FILE REFERENCE: PC-0050 US
 ; CURRENT APPLICATION NUMBER: US/09/907,372
 ; CURRENT FILING DATE: 2001-07-27
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 12
 ; LENGTH: 862
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; OTHER INFORMATION: Incyte ID No. US20020068242A1 8576918T1
 ; US-09-907-372-12

Query Match

Best Local Similarity 11.2%; Score 47.8; DB 10; Length 862;
 Best Local Similarity 74.5%; Pred. No. 1.8e-05;
 Matches 114; Conservative 0; Mismatches 32; Indels 7; Gaps 4;

QY 279 GTGGCCGACGATGACATGCTATGACCGGCGGCTCTATGACTATCTACGCAAC 338
 |||||||
 Db 853 GTGGCCGCGGTACCCAGGCAATAGGT-ACCGCGGTTCAAGACATACCGGCAAA 795

QY 339 A--TTTACATCTACATGAGCA--GTACTGGGGGAGACACCGGCTCTGGAGACCTCC 394
 || |||
 Db 794 AGTCGAATTCACACAGGAGCAAGTACTGGGGAGACACCGGCTCTGGAGACCTCC 735
 QY 395 AGCTACCCCGAAGC--TTCATACCCCATTTCCC 425
 || |||
 Db 734 AGCTACCCCGAAGCCTTCATACCCCATTTCCC 702

RESULT 15

US-09-844-864-22
 ; Sequence 22, Application US/09844864
 ; Patent No. US20020042926A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Matzduk, Martin
 ; APPLICANT: Ren, Yousheng
 ; APPLICANT: Wu, Xuemei
 ; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
 ; FILE REFERENCE: P01925052 / 09807797 / OIA 99-48
 ; CURRENT APPLICATION NUMBER: US/09/844,864
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 60/106,020
 ; PRIOR FILING DATE: 1998-10-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/25209
 ; PRIOR FILING DATE: 1999-10-28
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 809
 ; TYPE: DNA
 ; ORGANISM: mus musculus
 ; US-09-844-864-22

Query Match 8.3%; Score 35.4; DB 10; Length 809;
 Best Local Similarity 49.2%; Pred. No. 0.13;
 Matches 93; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 200 GGAGCGAGGGGTGCGGCAACAGCAGTCTGTGACCTGACAGGAGCGCAGTTGA 259
 || |||
 Db 610 GCAGCGCGGGTGGAGACAGCCACACCGAGAGACCGGAGACGTGTGGGGGATGCA 669
 QY 260 ACCCGGGGAGAGAGCCAGTGGCCACGCTACCAATGGCATTCATGTCACCGCGGCTC 319
 || |||
 Db 670 GTCTGAGCCTGGAGAGAGAGAGCAGTCTCTCCGACAGATGGCTCAGGACCCCGTGA 729
 QY 320 TATGACTATCTGCGCAACATCTACATCTACATGAGACGATCTGGGGGAGCCACCGGG 379
 || |||
 Db 730 TTCGATGCCCCCTCCCGCAAGACCACTATACCTTGACGAGCAAGAGCTCTCGCTT 789
 QY 380 TCCGTGAGA 388
 || |||
 Db 790 TCCAGGTGA 798

Search completed: April 15, 2003, 18:45:54
 Job time : 33.6131 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:39:00 ; Search time 100.94 Seconds
(without alignments)
526.717 Million cell updates/sec

Title: US-09-917-372-1
Perfect score: 2248
Sequence: 1 MLLPWATAPGLAWGLPLVIG.....EHCATPSNRGPRNQFIETHD 399

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :			
1:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:	*	
2:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:	*	
3:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:	*	
4:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:	*	
5:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:	*	
6:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:	*	
7:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:	*	
8:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:	*	
9:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:	*	
10:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:	*	
11:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:	*	
12:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:	*	
13:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:	*	
14:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:	*	
15:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:	*	
16:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:	*	
17:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:	*	
18:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:	*	
19:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:	*	
20:	/SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:	*	
21:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:	*	
22:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:	*	
23:	/SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:	*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1404	62.5	415	22 AAB36700	Human tumour necro
2	1325.5	59.0	305	23 ABE41926	Human ovarian anti
3	1129	50.2	197	18 AAW23220	Extracellular doma
4	1129	50.2	197	20 AAY31326	Human lymphotoxin
5	456	20.3	77	20 AAW94642	TNF-R extracellular
6	456	20.3	77	22 AAB69194	Human TNF-R extrac
7	363	16.1	461	16 AAR72504	p75 Tumour Necrosi
8	351.5	15.6	518	15 AAR51003	Sequence of a reco
9	349.5	15.5	461	12 AAR11001	40kD TNF inhibitor
10	349.5	15.5	461	12 AAR11141	Human TNF-R deduce

11	349.5	15.5	461	14 AAR42058	Fibroblast derived
12	349.5	15.5	461	21 AAB37801	Human tumour necro
13	349.5	15.5	461	21 AAB18717	A human tumour nec
14	349.5	15.5	461	21 AAB01342	Death receptor. H
15	349.5	15.5	461	22 AAB35331	Human TNF receptor
16	349.5	15.5	461	22 AAB36698	Human tumour necro
17	349.5	15.5	461	22 AAB37686	Human 40 kDa TNF i
18	349.5	15.5	461	23 AAU75172	Human TNF receptor
19	349.5	15.5	461	23 AAU75173	Human TNF receptor
20	349.5	15.5	461	23 AAU75174	Human TNF receptor
21	339.5	15.1	518	22 AAB70001	STNFR(075):Fc fusi
22	339.5	15.1	518	22 AAB50080	TNFR:Fc fusion pro
23	337.5	15.0	485	13 AAR24016	Fusion protein TNF
24	330	14.7	487	22 AAB50084	TNFR:Fc fusion pro
25	325	14.5	248	21 AAY94718	Human type 2 tumou
26	322.5	14.3	474	12 AAR11142	TNF-R deduced from
27	321.5	14.3	225	21 AAY77463	Primate protein se
28	321.5	14.3	227	22 AAB66981	Tnfr2 protein. Un
29	319	14.2	300	19 AAW66102	Amino acid sequenc
30	319	14.2	300	19 AAW63622	Human tumour necro
31	319	14.2	300	20 AAY03099	Human lung TNF-rec
32	319	14.2	300	20 AAY42182	Human FLINT #1 pro
33	319	14.2	300	20 AAY17479	Mammalian tumour n
34	319	14.2	300	20 AAY06817	Human DcR3 polyep
35	319	14.2	300	20 AAW97749	Human tumour necro
36	319	14.2	300	20 AAW95082	Orphan receptor (H
37	319	14.2	300	21 AAB19335	A full length huma
38	319	14.2	300	21 AAB28559	Human soluble TNF
39	319	14.2	300	21 AAB24057	Human PRO212 prote
40	319	14.2	300	21 AAB33416	Human PRO212 prote
41	319	14.2	300	21 AAB03621	Human Fas ligand i
42	319	14.2	300	21 AAY97246	M68 TNF receptor i
43	319	14.2	300	21 AAY90357	Human tumour necro
44	319	14.2	300	21 AAB24395	Human PRO212 prote
45	319	14.2	300	21 AAY96596	Human FLINT. Homo

ALIGNMENTS

RESULT 1
AAB36700
ID AAB36700 standard; Protein: 415 AA.
XX
AC AAB36700;
XX
DT 15-MAR-2001 (first entry)
XX
DE Human tumour necrosis factor receptor LTBR protein SEQ ID NO:6.
XX
KW Human; tumour necrosis factor receptor 5; TRID; TNFR-5; TR5; nontropic;
KW TRAIL receptor without intracellular domain; diagnosis; cytostatic;
KW tumour necrosis factor related apoptosis inducing ligand; vasotropic;
KW immunosuppressive; neuroprotective; antiviral; antiinflammatory;
KW anticonvulsant; antiparasitic; cardiant; anti-HIV; antiparkinsonian;
KW gene therapy; restenosis; graft versus host disease; tumour; cancer;
KW apoptotic cell death related disease; autoimmune disorder;
KW cardiovascular disorder; viral infection.
XX
OS Homo sapiens.
XX
PN WO200071150-A1.
XX
PD 30-NOV-2000.
XX
PF 18-MAY-2000; 2000WO-US13515.
XX
PR 20-MAY-1999; 990S-0135164.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Wei Y, Ruben SM, Gentz RL, Ni J;
XX

DR WPI; 2001-041051/05.

XX Nucleic acid encoding a TRID polypeptide, also referred to as tumor

PT necrosis factor receptor 5, useful in the diagnosis, treatment or

PT prevention of cancer, autoimmune disorders and viral infection.

XX Disclosure; Fig 2; 285pp; English.

XX The present invention describes the human TRID protein (tumour necrosis

CC factor (TNF) related apoptosis inducing ligand (TRAIL) receptor without

CC intracellular domain, also referred to as tumour necrosis factor

CC receptor 5 (TNR-5 or TR5). TRID has cytostatic, immunosuppressive,

CC notropic, neuroprotective, antiviral, antiinflammatory, anticonvulsant,

CC antiparasitic, cardiant, anti-HIV, antiparkinsonian and vasotropic

CC activities, and can be used in gene therapy. The TRID polynucleotides

CC are useful for detecting complementary polynucleotides. TRID proteins and

CC polynucleotides are useful in the treatment of tumours, resistance to

CC parasite, bacteria and viruses, restenosis and graft versus host disease.

CC They are also useful for inducing proliferation of T-cells, endothelial

CC cells and certain haematopoietic cells, to regulate antiviral responses

CC and to prevent certain autoimmune diseases after stimulation of TRID by

CC an agonist or TRAIL binding facilitator. The antibodies which bind TRID

CC polypeptides are useful for treating and/or preventing diseases

CC associated with increased or decreased apoptotic cell death. The TRID

CC polynucleotides, proteins, antibodies, agonists and antagonists are

CC useful in the diagnosis, treatment or prevention of: (a) cancer;

CC (b) autoimmune disorders; (c) diseases associated with increased

CC apoptosis; (d) cardiovascular disorders; and (e) viral infection. The

CC present sequence represents a tumour necrosis factor receptor used in

CC comparison with TRID in the exemplification of the present invention.

XX

SQ Sequence 415 AA;

Query Match 62.5%; Score 1404; DB 22; Length 415;

Best Local Similarity 63.9%; Pred. No. 3.6e-85;

Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

QY 1 MLLPWATSAAGLWGLVGLFLLAASQAPVPPYASENQTCDQKEYEYEPQHRCCS 60

Db 1 MRLPRASSPCGLWGLVGLFLLAASQAPVPPYASENQTCDQKEYEYEPQHRCCS 60

QY 61 RCPPTGYYSKSRRTDVCATCAENSNEHWNLYTICQLCPDPPVGLBEIACTSKR 120

Db 61 RCPPTGEFFVAVCSQDVTCTCPHNSNEHWNHLSLTCQLCPDPPVGLBEIACTSKR 120

QY 121 KTCRCQCPGMFCAAWALECTHC--ELLSDCPPTGAELKDEVGKGNHCVPCAKGHFQNT 178

Db 121 KACRCQCPGMFCAAWALECTHC--ELLSDCPPTGAELKDEVGKGNHCVPCAKGHFQNT 180

QY 179 SSPSARCQPHTRCENQGLVEAAGTAQSDTTCKNPLEP----- 216

Db 181 SSPSARCQPHTRCENQGLVEAAGTAQSDTTCKNPLEP----- 240

QY 217 -----LPPEMS---GSLKRRPQGGPNPAGVPPKAHPYFDLPVLPISGDVSPV 268

Db 241 ACAMRHPSLCKLTKLRHPEGESPPCPA---PRADPHFPDLAEPPLPMSGDLSPS 296

QY 269 STGLPAAPVLEAGVQQOSPLDITREPQLERQSOVAHGNGIHVTGSMITGNIIY 328

Db 297 PAGPTAPSLSEVLVQQOSPLVQARELEAEPGEHQVAHGNGIHVTGSMITGNIIY 356

QY 329 NGPVLGGPPGGLDPATPEPPYPIPEEGDPGPGPLSTPHQEDGKAHLAEYEHCG 383

Db 357 NGPVLGGTGFCDPPAPPEPPYPTPEGAPGSELSTPYQEDGKAHLAEYETLG 411

RESULT 2

ABP41926

ID ABP41926 standard; Protein; 305 AA.

XX

AC ABP41926;

XX

DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HSAJ44, SEQ ID NO:3058.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

XX KW ovarian cancer; breast cancer; tumour; reproductive system disorder;

XX KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

XX KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;

XX KW inflammatory condition; immune disorder; blood disorder;

XX KW cardiovascular disorder; respiratory disorder; neurological disorder;

XX KW gastrointestinal disorder; urinary system disorder; drug screening;

XX KW gene therapy; chromosome mapping; forensic analysis;

XX KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;

XX KW antiinflammatory; gynaecological; reproductive; chromosome 12p13.

XX Homo sapiens.

OS WO200200677-A1.

PN 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US18569.

XX 07-JUN-2000; 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

PI WPI; 2002-147878/19.

XX N-PSDB; ABQ53003.

DR

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,

PT useful in the prevention, treatment and diagnosis of cancer (e.g.

PT ovarian cancer), immune disorders, cardiovascular disorders and

PT neurological diseases.

XX Claim 11; SEQ ID NO 3058; 2922pp; English.

PS The invention relates to 2175 novel human ovarian antigens (ABP41054-

XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also

CC encompasses polypeptides 90% identical and polynucleotides 95% identical

CC to the sequences of the invention. The invention additionally relates to

CC recombinant vectors and host cells comprising human ovarian antigen

CC polynucleotides, antibodies against human ovarian antigens, and the use

CC of ovarian antigen polynucleotides and polypeptides in diagnosing,

CC treating, prognosing or preventing various ovary and/or breast-related

CC disorders. Such conditions include ovarian cancer and breast cancer, and

CC metastatic tumours of ovarian or breast origin, reproductive system

CC disorders (e.g., infertility, disorders of pregnancy, anovulation,

CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine

CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic

CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and

CC vaginitis), immune disorders (e.g., congenital and acquired

CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),

CC blood-related disorders (e.g., anaemia), cardiovascular disorders,

CC respiratory disorders, neurological disorders, gastrointestinal disorders

CC and urinary system disorders. Ovarian antigen polypeptides and

CC polynucleotides may also be used in screening for compounds which

CC modulate ovarian antigen expression or activity. The polynucleotides may

CC further be used for gene therapy, chromosome mapping, in the

CC identification of individuals and in forensic analysis, and the

CC polypeptides may be used as food additives or to prepare antibodies

CC useful in disease diagnosis, drug targeting and phenotyping. The present

CC sequence represents a human ovarian antigen of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 305 AA;

Query Match 59.0%; Score 1325.5; DB 23; Length 305;

Best Local Similarity 85.7%; Pred. No. 4e-80;

Matches 246; Conservative 0; Mismatches 4; Indels 37; Gaps 2;

CC inhibitor of the interaction between LT-beta and its receptor can be
 CC administered for altering the survival or maintenance of follicular
 CC dendritic cells in a subject and for altering the architecture of the
 CC organs of the immune system. The method is useful for treating tumours,
 CC specifically follicular lymphomas. It offers an alternative therapy for
 CC those with tumours resistant to traditional chemotherapy. The present
 CC sequence represents the extracellular region of the human LTbeta-receptor
 CC and comprises the ligand binding domain.

SQ Sequence 197 AA;
 Query Match 50.2%; Score 1129; DB 20; Length 197;
 Best Local Similarity 99.5%; Pred. No. 2.5e-67;
 Matches 196; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 28 SQQAVPPYASENQCRDQKEYEYEPQHRICCSRCPPGTYYSAKSRIRDTVCATCAENS 87
 DB 1 SQQAVPPYASENQCRDQKEYEYEPQHRICCSRCPPGTYYSAKSRIRDTVCATCAENS 60
 QY 88 YNEHNNYLTICQLCRPCDPVVMGLEIAPCTSKRKTQCRCPGMFCAAWALECTHCELLSD 147
 DB 61 YNEHNNYLTICQLCRPCDPVVMGLEIAPCTSKRKTQCRCPGMFCAAWALECTHCELLSD 120
 QY 148 CPPGTEAEALKDEVGKGNHCVCKAGHFQNTSSPSARCPQHTRCENQGLVEAAPGTAQSD 207
 DB 121 CPPGTEAEALKDEVGKGNHCVCKAGHFQNTSSPSARCPQHTRCENQGLVEAAPGTAQSD 180
 QY 208 TTCKNPLEPLPPMSGS 224
 DB 181 TTCKNPLEPLPPMSGT 197

RESULT 5
 AAW94642
 ID AAW94642 standard; peptide; 77 AA.

XX AC AAW94642;
 DT 29-APR-1999 (first entry)
 XX TNF-R extracellular Cys-rich domain TNF-R-rp.
 XX Tumour necrosis factor receptor; TNF-R; autoimmune diseases;
 KW inflammation; septic shock; cachexia; graft versus host disease;
 KW skin allergic reaction; immune complex disease; malaria;
 KW transplantation rejection.
 XX Homo sapiens.
 OS
 XX WO9853842-A1.
 PN
 PD 03-DEC-1998.
 PF 29-MAY-1998; 98WO-US10891.
 XX
 PR 30-MAY-1997; 97US-0866545.
 XX (UYPE-) UNIV PENNSYLVANIA.
 PA
 PI Greene MI, Murali R, Takasaki W;
 XX WPI; 1999-080781/07.
 DR
 XX New compounds designed from a binding loop of a tumour necrosis
 PT factor receptor - are capable of inhibiting the biological
 PT activities of tumour necrosis factor, e.g., in treating inflammation
 PT or autoimmune diseases
 XX
 PS Disclosure; Fig 1; 78pp; English.
 XX
 PS The present invention describes peptides and peptide analogues which
 CC correspond in primary sequence to a binding loop of a tumour necrosis
 CC factor receptor (TNF-R) superfamily member. The compounds are especially

CC designed from a binding loop of TNF-R p55. They are capable of
 CC inhibiting TNF binding to its cellular receptors and may be used to
 CC inhibit the biological activities of TNF. They may be used in treating
 CC TNF-associated conditions such as acute and chronic inflammatory
 CC responses, septic shock, cachexia, autoimmunity, graft-versus-host
 CC disease, skin allergic reactions, immune complex disease,
 CC transplantation rejection and malaria. Administration is, e.g. oral,
 CC transdermal, transmucosal, pulmonary, subcutaneous, intravenous or
 CC intramuscular. Parenteral dosage is 0.1-5 mg/kg/day. The present
 CC sequence represents an extracellular Cys-rich domain of TNF-R from the
 CC present invention.

SQ Sequence 77 AA;
 Query Match 20.3%; Score 456; DB 20; Length 77;
 Best Local Similarity 100.0%; Pred. No. 3.2e-23;
 Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 79 VCATCAENSYNHNNYLTICQLCRPCDPVVMGLEIAPCTSKRKTQCRCPGMFCAAWALE 138
 DB 1 VCATCAENSYNHNNYLTICQLCRPCDPVVMGLEIAPCTSKRKTQCRCPGMFCAAWALE 60
 QY 139 CTHCELLSDCPPGTEAE 155
 DB 61 CTHCELLSDCPPGTEAE 77

RESULT 6
 AAB69194
 ID AAB69194 standard; protein; 77 AA.

XX AC AAB69194;
 DT 30-APR-2001 (first entry)
 XX Human TNF-R extracellular Cys-rich domain TNF-R-rp SEQ ID NO:3.
 DE Tumour necrosis factor receptor; TNF-R; inhibition; osteoclast;
 XX osteoclastogenesis; bone loss; bone resorption; osteopathic; cytostatic;
 KW antirheumatic; antiarthritic; antiinflammatory; immunomodulatory;
 KW tumour necrosis factor-related activation-induced cytokine; TRANCE;
 KW receptor activator of NF-kappaB ligand; RANK; osteoporosis;
 KW Paget's disease; metastatic bone disease; rheumatoid arthritis;
 KW periodontal disease; modulating dendritic cell maturation;
 KW T cell proliferation; CD40 receptor system.
 XX Homo sapiens.
 OS
 XX WO200108699-A1.
 PN
 PD 08-FEB-2001.
 PF 28-JUL-2000; 2000WO-US20510.
 XX
 PR 28-JUL-1999; 99US-0146090.
 XX (UYPE-) UNIV PENNSYLVANIA.
 PA (AOKI/) AOKI K.
 PA (HORN/) HORNE W C.
 PA (BARO/) BARON R.
 XX
 PI Aoki K, Horne WC, Baron R, Greene MI, Murali R;
 XX WPI; 2001-182866/18.
 DR
 XX Use of peptides and peptide analogs which are TRANCE/RANK inhibitors,
 PT for inhibiting osteoclastogenesis and bone resorption
 PT
 XX Disclosure; Fig 1; 81pp; English.
 PS
 XX The present invention describes a method for inhibiting
 CC osteoclastogenesis and bone resorption. Osteoclastogenesis and bone
 CC resorption inhibiting peptide analogues from the present invention have

CC osteopathic, cytostatic, antirheumatic, antiarthritic, antiinflammatory
 CC and immunomodulatory activities, and are tumour necrosis factor (TNF)-
 CC related activation-induced cytokine (TRANCE)/ receptor activator of
 CC NF-kappaB ligand (RANK) inhibitors. The method is useful for treating
 CC diseases characterised by bone loss such as osteoporosis, Paget's
 CC disease, metastatic bone disease, rheumatoid arthritis or periodontal
 CC disease, and modulating dendritic cell maturation, T cell proliferation,
 CC and/or CD40 receptor systems. The present sequence represents an
 CC extracellular Cys-rich domain of a tumour necrosis factor receptor
 CC (TNF-R) superfamily member, which is used in the exemplification of
 CC the present invention.

XX Sequence 77 AA;

Query Match 20.3%; Score 456; DB 22; Length 77;

Best Local Similarity 100.0%; Pred. No. 3.2e-23;

Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 VCATCAENSYNEHWNLYTICQLCRPDVPMGLEIAPCTSKRTQCRQCPGMFCAAWALE 138

Db 1 VCATCAENSYNEHWNLYTICQLCRPDVPMGLEIAPCTSKRTQCRQCPGMFCAAWALE 60

QY 139 CTHCELLSDCPGTEAE 155

Db 61 CTHCELLSDCPGTEAE 77

RESULT 7

AAR72504

ID AAR72504 standard; Protein; 461 AA.

XX AC

XX AAR72504;

XX 31-OCT-1995 (first entry)

XX p75 Tumour Necrosis Factor Receptor.

DE DE

XX KW

Ligand; tumour necrosis factor; nerve growth factor; TNF; NGF;

XX KW

XX OS

Homo sapiens.

XX FH

Key Location/Qualifiers

FT Region 27..214

FT /label= TBPII.

FT Domain 258..285

FT /label= Transmembrane domain.

FT Misc-difference 259

FT /note= "Unidentified amino acid."

XX EP648783-A.

XX PN

XX 19-APR-1995.

PD PD

XX XX

XX 11-OCT-1994; 94EP-0116015.

XX PF

XX 12-OCT-1993; 93IL-0107267.

XX PR

XX (YEDA) YEDA RES & DEV CO LTD.

PA PA

PA (WALL/) WALLACH D.

XX XX

XX Beletsky I, Bigda J, Mett I, Wallach D;

XX XX

XX WPI; 1995-148673/20.

DR DR

XX N-PSDB; AAQ89544.

XX XX

XX Tumour necrosis factor (TNF) receptor ligand - used to increase

PT PT

XX Inhibitory effect of a soluble TNF receptor

XX XX

PS Disclosure; Figure 2; 18pp; English.

XX XX

XX A ligand to a member of the tumour necrosis factor (TNF)/nerve

CC growth factor (NGF) receptor family which binds either to the region

CC of the 4th-Cys rich domain of the receptor, or to the region between
 CC it and the cell membrane may be used in the production of a
 CC pharmaceutical composition for increasing the inhibitory effect of a
 CC soluble receptor of the TNF/NGF receptor family. This sequence
 CC is the sequence of the p75 TNF receptor.

XX Sequence 461 AA;

Query Match 16.1%; Score 363; DB 16; Length 461;

Best Local Similarity 29.1%; Pred. No. 3.5e-16;

Matches 119; Conservative 49; Mismatches 159; Indels 82; Gaps 23;

QY 9 APLGAWPLVLGLFGLLAASQPAVP-----PYASE-NOTCRDQERKEYEPOHRCISR 61

Db 2 APVAVMAALAVGL-ELWAAA--HALPAQVAFTPYAPEPGSTCR--LREYDQTAQMCCSK 56

QY 62 CPFGTYYSAKSRIRDFVCATCAENSYNEHWNLYTICQLCR---PCDPVMGLEIAPCTS 118

Db 57 CSPGQHAFAFCTKTSDFVCDSCEDSTYQLMNWVPECLSCGRSSDOV----ETQACTR 112

QY 119 KRKTQCRQCPGMFCAAWALE-CTHCELLSDCPGQ-----TEAEKLDVKGKNNHCVPCA 172

Db 113 EQNRICTRFGWCALSKQEGCRLCAELKRCRFGFGVARGTETSDV-----CKFCAP 166

QY 173 GHFQNTSSPSARCOPHTRCENOGIVEAAPGTAQSDTTC--KNPLELPPEMSGSL----- 225

Db 167 GTFSTTSSTDICRPHQICN---VVAIPGNASMDAVCTSTSPTRSMAP---GAVHLPQP 219

QY 226 LKRRPQEGPNVAGSWEP---PKAHYFPDLVQPLLPISGDVS----PVS--TGLPAAP 276

Db 220 VSTRSHTQPTP-----EPSTAPSTSLFP--MGPSPPAEGSTGDFLPLVGLVGVVATLG 272

QY 277 VLEAGVPQ-----QQSPDLDTREPOLE--PGEOSOVAHGTNGIHV-----TGSMT 320

Db 273 LLIGVVNCVIMTQVKKKPLCLQREAKVPHLPADKARGTQGPQQHLLITAPSSSSSLE 332

QY 321 ITGNIYINGVLGGPPGPDLPATPEPPYPIPEEGDPGPPGLSTPHQE 369

Db 333 SSASALDRRAPTRNQPOGAGEARASTG-----SSDSSPGHGHTQAPGVE 376

RESULT 8

AAR51003

ID AAR51003 standard; Protein; 518 AA.

XX AC

XX AAR51003;

XX 07-OCT-1994 (first entry)

DT DT

XX Sequence of a recombinant human (rhu) tumour necrosis factor receptor

DE TNFR/fc fusion protein.

XX XX

XX Tumour necrosis factor receptor; chimeric antibody molecule;

XX KW

XX Immuglobulin.

XX OS

XX Synthetic.

XX PN

XX WO9406476-A.

XX XX

XX 31-MAR-1994.

XX PF

XX 14-SEP-1993; 93WO-US08666.

XX XX

XX 15-SEP-1992; 92US-0946236.

XX PR

XX (IMMV) IMMUNEX CORP.

XX PA

XX Jacobs CA, Smith CA;

XX PI

XX WPI; 1994-118172/14.

XX DR

XX N-PSDB; AAQ45225.

XX XX

XX Treating TNF mediated inflammatory diseases with TNF antagonist -

PT PT

PT esp. soluble form of TNF receptor, opt. as fusion protein with
 PT human immunoglobulin Fc region, esp. for treating arthritis
 XX Disclosure; Page 32-34; 47pp; English.
 PS
 XX AAQ45224 is cDNA from clone 1 of library WI-26 V44 of human fibroblast
 CC cell line WI-26 V44. The mature full-length TNFRI is a glycoprotein
 CC having a mol. wt. of about 75-80 kDa. The cloning of the cDNA for
 CC TNFRI was described in Smith et al., Science 248:1019, 1990. Clone 1
 CC is contained in expression vector pCAV/NOT-TNFR (ATCC 68088). A
 CC recombinant chimeric antibody may be produced having TNFR sequences
 CC substituted for the variable domains of either or both of the
 CC immunoglobulin molecule heavy and light chains and having
 CC unmodified constant region domains. A specific example of a TNFR/Fc
 CC fusion protein is given in AAQ45225/R51003. The rhu TNFR/Fc fusion
 CC gene was created by ligating the following fragments into a cloning
 CC vector: 1) an 867 bp Asp718-Pvu2 fragment from pCAV/NOT-TNFR (ATCC
 CC 68088) contg. the cDNA encoding the truncated TNFR. 2) a 700 bp
 CC Styl-Spel fragment from plasmid pIXY498 coding for 232 AAs of the
 CC Fc portion of human IgG1. Plasmid pIXY498 is a yeast expression
 CC vector contg. the Fc fragment of human IgG1. 3) An oligo linker, to
 CC fuse the truncated TNFR with the human IgG1 Fc fragment. This
 CC linker was created by PCR using primer AAQ45226, which encodes the 3'
 CC end of the truncated TNF receptor and the 5' end of human IgG1,
 CC and primer AAQ45227, which is an antisense sequence encoding bps
 CC 257-237 of human IgG1.
 XX
 SQ Sequence 518 AA;
 Query Match 15.6%; Score 351.5; DB 15; Length 518;
 Best Local Similarity 27.6%; Pred. No. 2.3e-15;
 Matches 106; Conservative 32; Mismatches 127; Indels 119; Gaps 16;
 Qy 4 PWATSAAGLPLVLGI---FGLIAASQPOAVPYASE-NOTCRDQKEYEYPOHRIICC 59
 Db 26 PPASMAPVAVWAALAVGLWAGHALPAQVAFYAPPEPGSTCR--LREYDQTAQMCCK 83
 Qy 60 SRCPPGTYSKRSIRDTVCATCAENSYNEHWNLYTICQLCR---PCDPVMGLEEIAPC 116
 Db 84 SKCSPQAHKVFCTKTSDTVCDSCEDSTYTLWNVWPECLSCGSRCSQDV---ETQAC 139
 Qy 117 TSKRKTQRCQPGMFCAMWALE-CTHCELLSDCPPG-----TEAEKDEVGKGNHVCVPC 170
 Db 140 TREQNRICTRPGWYCALSKQEGRCRLCAPLRKCRPGFVGARPGTETSDVW-----CKPC 193
 Qy 171 KAGHPONTSSPSARCOPHTRCENQGLVEAAPGTAQSDTTCNPLPLPPMSSGSLKKRP 230
 Db 194 APGTSNTSSTDICRPHQICN-----VVAIPGNASMDAVCTIS-----231
 Qy 231 QGEGNPVAGSWEPKPAHYFPDLVQPLLPISGDVSPYSTGLPAAPVLEAGVPQQSPLD 290
 Db 232 -----TSPTSRMAPGA-----VHLPPQVPS---250
 Qy 291 LTREPQLEGGESQVAHGHTNGHVTGSMITIGNIYIYINGPVLGGPPGDLPATPE---347
 Db 251 -TRSQHTQPTPEPTAPSTFLLPMGPS-----PPAEGSTGDEPKSCD 292
 Qy 348 -----PPYPIPE-EGDPG-----pp 361
 Db 293 KHTHTCPCPAPELGGPSVFLFPP 316
 RESULT 9
 AAR11001
 ID AAR11001 standard; Protein; 461 AA.
 XX AAR11001;
 AC AAR11001;
 XX
 DT 13-MAY-1991 (first entry)
 XX
 DE 40kD TNF inhibitor precursor.
 XX
 KW Tumour necrosis factor; inhibitor.

XX Homo sapiens.
 OS
 XX AU058976-A.
 PN
 XX 24-JAN-1991.
 PD
 XX 16-JUL-1990; 90AU-0058976.
 PF
 XX 07-FEB-1990; 90US-0479661.
 PR 18-JUL-1989; 89US-0381080.
 PR 11-DEC-1989; 89US-0450329.
 XX
 PA (SYNE-) SYNERGEN INC.
 XX
 XX WPI; 1991-073847/11.
 DR N-PSDB; AAQ10907.
 XX
 XX Tumour necrosis factor inhibitor - for suppression of TNF-alpha
 PT and -beta, useful as therapeutic agent.
 PS Disclosure; Fig 39; 142pp; English.
 XX
 CC The sequence comprises the entire 40 kD TNF inhibitor. The clone
 CC from which the sequence was deduced was isolated from a cDNA
 CC library prepd. from RNA form U937 cells treated with PMA/PHA.
 CC The whole gene can be inserted into expression vectors for prepn.
 CC of TNF inhibitor for use in the treatment of inflammatory and
 CC degenerative diseases.
 CC See also AAR10986 and AAR10984.
 XX
 SQ Sequence 461 AA;
 Query Match 15.5%; Score 349.5; DB 12; Length 461;
 Best Local Similarity 29.7%; Pred. No. 2.7e-15;
 Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;
 Qy 9 APGLAWPLVLGLFLLAASQPOAVP-----PYASE-NOTCRDQKEYEYPOHRIICSR 61
 Db 2 APVAVWAALAVGL-ELWAAA--HALPAQVAFYAPPEPGSTCR--LREYDQTAQMCCK 56
 Qy 62 CPPGTYSKRSIRDTVCATCAENSYNEHWNLYTICQLCR---PCDPVMGLEEIAPCS 118
 Db 57 CSPQAHKVFCTKTSDTVCDSCEDSTYTLWNVWPECLSCGSRCSQDV-----ETQACTR 112
 Qy 119 KRKTQRCQPGMFCAMWALE-CTHCELLSDCPPG-----TEAEKDEVGKGNHVCVPCA 172
 Db 113 EQNRICTRPGWYCALSKQEGRCRLCAPLRKCRPGFVGARPGTETSDVW-----CKCAP 166
 Qy 173 GHFONTSPPSARCOPHTRCENQGLVEAAPGTAQSDTTC--KNPLEPLPPMSSGSL-----225
 Db 167 GTFSTNTSSTDICRPHQICN-----VVAIPGNASMDAVCTISPTSRMAP---GAVHLPQP 219
 Qy 226 LKRPOGPNVAGSWEPKPAH-----PYFPDLVQPLLPISGDVSPVS--TGLPAAP 276
 Db 220 VTRSQHTQPTPEPTAPSTFLLPMGPS--PPAEGSTGDFALPVLGLVGVVTAIG 272
 Qy 277 VLEAGVPQ-----QOSPDLTREPQLE--PGEQSQVAHGHTNGIHV-----TGGSM 320
 Db 273 LLIGVNCVIMTQVKKKPLCLQREAKVPHLPADKARTQGEQHQHLLITAPSSSSSLE 332
 Qy 321 ITGNIYIYINGPVLGGPPGPG 340
 Db 333 SSASALDRRAPTRNQPOAPG 352
 RESULT 10
 AAR11141
 ID AAR11141 standard; Protein; 461 AA.
 XX AAR11141;
 AC AAR11141;
 XX
 DT 24-MAY-1991 (first entry)

AC AAB18717;
XX
DT 22-JAN-2001 (first entry)
XX
DE A human tumour necrosis factor family receptor (TNF-RII).
XX
KW Human; tumour necrosis factor family receptor; TRI; tumour growth;
KW cell proliferation; chlamydia infection; immunodeficiency; septic shock;
KW T-cell mediated autoimmune disease; acquired immunodeficiency syndrome;
KW AIDS; cerebral malaria; graft rejection; cytotoxicity; cachexia;
KW apoptosis; inflammation; cancer; cardiovascular disease; angiogenesis;
KW inflammatory disease; atherosclerosis; diabetes mellitus; allergy;
KW neurological disorder; autoimmune disease; wound healing; bone formation;
KW osteoporosis.
XX
OS Homo sapiens.
XX
PN WO200054651-A2.
XX
PD 21-SEP-2000.
XX
PF 15-MAR-2000; 2000WO-US06592.
XX
PR 15-MAR-1999; 99US-0124489.
PR 26-MAY-1999; 99US-0136248.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Greene JM, Fleischmann RD, Ni J;
XX
DR WPI; 2000-618858/59.
XX
XX Novel tumour necrosis factor family receptor for diagnosing and treating
PT acquired immunodeficiency syndrome, cancer, cardiovascular diseases,
PT inflammatory diseases and autoimmune diseases
XX
PS Claim 16; Page 24-25; 228pp; English.
XX
XX The present sequence encodes human tumour necrosis factor family
CC receptor (TNF-RII) polypeptide. The specification describes a TNF
CC receptor designated TRI. An agonist to the TRI receptor is useful for
CC inhibiting tumour growth, to stimulate human cellular proliferation, to
CC regulate immune response and antiviral response, to protect against the
CC effects of ionising radiations, to protect against chlamydia infections,
CC to regulate growth, and to treat immunodeficiencies such as in human
CC immunodeficiency virus (HIV). An antagonist to the TRI receptor is
CC useful for treating T-cell mediated autoimmune diseases, acquired
CC immunodeficiency syndrome (AIDS), septic shock, cerebral malaria, graft
CC rejection, cytotoxicity, cachexia, apoptosis, and inflammation. TRI
CC polynucleotides and polypeptides, and TRI agonists and antagonists
CC are useful for treating cancers, cardiovascular diseases, inflammatory
CC diseases, atherosclerosis, diabetes mellitus, neurological disorders,
CC autoimmune diseases, for promoting angiogenesis, for treating allergy,
CC for wound healing, for regulating bone formation and for treating
CC osteoporosis.
XX
SQ Sequence 461 AA;
Query Match 15.5%; Score 349.5; DB 21; Length 461;
Best Local Similarity 29.7%; Pred. No. 2.7e-15;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;
QY 9 APGLAWGPLVLGLFLLAASOPQAVP-----PYASE-NQTCRQDEKEYEYEPQHRICCSR 61
DB 2 APVAVWAALAVGL-ELWAAA--HALPAQVATPYAPEGSTCR--LREYYDQTAQMCCSK 56
QY 62 CPPTGYVSAKCSRRIDTCATCAENSNEHWNLYTLICQLCR---PCDPVMGLEELIAPCTS 118
DB 57 CSPGQHAQVFCITKTSDDTVDCEDSTYQLNWNVPECLSCGRSSDQV---EFQACTR 112
QY 119 KRKTCRCQPGMFCMAWALE-CTHCELLSDCPPG-----TRAEKDEYKGNHNCVPCKA 172
DB 113 EQNRICTRPGWYCALSKQEGRCGLCAPLRKCRPGFVARPGTGTSDVV-----CKPCAP 166

QY 173 GHFONTSPPSARCOPHTRCENOGVLVEAAPGTAQSDTTC--KNPLEPLPEMGSGL----- 225
DB 167 GTFSNTTSTTDICRPHQICN-----VVAIPGNASMDAVCTSTSPTRSMAP---GAVHLPQP 219
QY 226 LKRRPQEGEPNPVAGSWEPKKAH-----PYFPDLVQPLLPISGDVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTSELLPMGPS--PPAEGSTGDFALPVGLIVGTALG 272
QY 277 VLEAGVQ-----QQSPDLTREPQLE--PGEOSQVAHGTNGIHV-----TGGSM 320
DB 273 LLIIIGVNCVIMTQVKKPLCLQREAKVPHLPADKARGTQGPQQHLLITAPSSSSSLE 332
QY 321 ITGNIYIYNGPVLGPPGPG 340
DB 333 SSASALDRRAPTRNQOPAG 352
RESULT 14
AAB01342
ID AAB01342 standard; Protein; 461 AA.
XX
AC AAB01342;
XX
DT 25-SEP-2000 (first entry)
XX
DE Death receptor.
XX
KW ULI44; death receptor; apoptosis; programmed cell death; FAS;
KW TNF-R1; TRAMP; DR-6; TRAIL; modulation; treatment; cancer; virus;
KW human.
XX
OS Homo sapiens.
XX
PN WO2000034335-A2.
XX
PD 15-JUN-2000.
XX
PF 03-DEC-1999; 99WO-US26035;
XX
PR 04-DEC-1998; 98US-0205018.
XX
PA (SCHE) SCHERING CORP.
XX
PI Leong C, Phillips JH;
XX
DR WPI; 2000-423383/36.
XX
XX Purified or recombinant polypeptide for modulating apoptosis comprises
PT a sequence which binds to an antibody specific for ULI44 or its
PT fragments
XX
PS Disclosure; Page 74-75; 76pp; English.
XX
XX A pure or recombinant polypeptide which binds to a polyclonal antibody
CC specific for the mature ULI44 is useful for screening molecules which
CC block induction of apoptosis or interfere with antiapoptotic activity.
CC The polypeptide is also useful for modulating apoptosis and useful in
CC treatment of conditions associated with abnormal physiology or
CC regulation, such as cancer or degenerative conditions and for
CC regulation of viral infection and replication. At least five
CC different death receptors are known, which include the CD95
CC (Fas/APO-1), the TNF receptor-1, TNF receptor apoptosis-mediated
CC protein (TRAMP), death receptor-6 (DR-6), and TNF-related
CC apoptosis-inducing ligand (TRAIL) receptors 1, 2 and 4.
XX
SQ Sequence 461 AA;
Query Match 15.5%; Score 349.5; DB 21; Length 461;
Best Local Similarity 29.7%; Pred. No. 2.7e-15;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;
QY 9 APGLAWGPLVLGLFLLAASOPQAVP-----PYASE-NQTCRQDEKEYEYEPQHRICCSR 61

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:45:16 ; Search time 102.845 Seconds
(without alignments)
799.386 Million cell updates/sec

Title: US-09-917-372-1

Perfect score: 2248

Sequence: 1 MLLPWATSAPLGLVGLG.....EHGATPSNRGPRNQFIITHD 399

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phase.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertibrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriap.*
- 17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	340	15.1	433	11 Q912M6	Q912M6 rattus norv
2	312	13.9	459	11 Q62327	Q62327 mus musculus
3	311.5	13.9	482	11 Q88734	Q88734 mus musculus
4	303.5	13.5	278	6 Q8SQ34	Q8sq34 sus scrofa
5	271	12.1	277	6 Q8WMO2	Q8wmq2 ovis aries
6	266	11.8	651	13 Q98SM6	Q98sm6 gallus gall
7	253	11.3	276	13 Q9DD2	Q9ddd2 gallus gall
8	252.5	11.2	267	6 Q02764	Q02764 oryctolagus
9	248	11.0	186	12 Q72735	Q72735 cowpox viru
10	244	10.9	283	6 Q9XSZ8	Q9xsx8 cercopithec
11	242.5	10.8	302	13 Q9PUS0	Q9pus0 salvelinus
12	236	10.5	285	13 Q9OW71	Q9ow71 oncorhynch
13	233	10.4	186	12 Q9FP87	Q9fp87 cowpox viru
14	230	10.2	351	12 Q75117	Q75117 cowpox viru
15	228	10.1	285	13 Q90YS6	Q90ys6 oncorhynch
16	227	10.1	186	12 Q911R5	Q911r5 vaccinia vi

17	226.5	10.1	167	12 Q8UYL3	Q8uyl3 vaccinia vi
18	225	10.0	349	12 Q57098	Q57098 camelpox vi
19	225	10.0	349	12 Q57099	Q57099 monkeypox v
20	225	10.0	349	12 Q8UYA7	Q8uya7 camelpox vi
21	225	10.0	349	12 Q57284	Q57284 camelpox vi
22	223.5	9.9	446	6 Q95ND3	Q95nd3 felis silve
23	222.5	9.9	169	11 Q9JKE0	Q9jke0 rattus norv
24	222	9.9	186	12 Q9WJB4	Q9wjb4 vaccinia vi
25	221	9.8	349	12 Q57097	Q57097 camelpox vi
26	221	9.8	349	12 Q57100	Q57100 monkeypox v
27	219	9.7	349	12 Q57102	Q57102 monkeypox v
28	219	9.7	349	12 Q57291	Q57291 monkeypox v
29	218.5	9.7	167	12 Q9DJL2	Q9djl2 cowpox viru
30	218.5	9.7	350	12 Q57116	Q57116 cowpox viru
31	218	9.7	348	12 Q57103	Q57103 monkeypox v
32	218	9.7	348	12 Q57108	Q57108 monkeypox v
33	216.5	9.6	167	12 Q72762	Q72762 cowpox viru
34	216	9.6	348	12 Q57277	Q57277 monkeypox v
35	216	9.6	349	12 Q57101	Q57101 monkeypox v
36	210	9.3	348	12 Q57112	Q57112 variola vir
37	210	9.3	348	12 Q85407	Q85407 variola vir
38	208	9.3	351	12 Q57121	Q57121 cowpox viru
39	207	9.2	349	12 Q57110	Q57110 variola vir
40	207	9.2	349	12 Q57111	Q57111 variola vir
41	207	9.2	349	12 Q89098	Q89098 variola vir
42	207	9.2	349	12 Q89118	Q89118 variola vir
43	207	9.2	351	12 Q73559	Q73559 cowpox viru
44	205	9.1	347	12 Q57119	Q57119 cowpox viru
45	203.5	9.1	328	6 Q9BDP0	Q9bdp0 aotus trivi

ALIGNMENTS

RESULT 1

Q912M6 PRELIMINARY; PRT; 433 AA.
ID Q912M6;
AC Q912M6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Tumor necrosis factor receptor type II (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Osburg B., Peiser C., Doemling D., Schomburg L., Voigt K., Bickel U.;
RT "TNF-receptors p60 and p80 are constitutively expressed by rat brain
capillary endothelial cells and participate in TNF-alpha transport
through the blood-brain barrier.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF420214; AAL16021.1; -;
DR InterPro: IPR001368; TNFR_c6.
DR Pfam: PF00020; TNFR_c6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_2.
DR PROSITE; PS00505; TNFR_NGFR_2; 3.
KW Receptor.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45723 MW; 75736D835E72CA4A CRC64;

Query Match 15.1%; Score 340; DB 11; Length 433;

Best Local Similarity 30.0%; Pred. No. 2.9e-19;

Matches 104; Conservative 45; Mismatches 144; Indels 54; Gaps 19;

QY 35 PYASENQTCDQKEYPEPHRICCRCPPTGVSAKCSRIKRTVCATCAENSNEHWNY 94

DB 11 PYKPEPNQCGQISQYDKKQAKCCAKCPGPQYAKHKTSTVCAACGAAGMFTQWNH 70

QY 95 LTICQLC-RPC--DPVNGLEIAPCTSKRKTCRCQPGMFC--AWALECTHCELLSDCP 149

```

Db 71 LHTCLSCSSCDDQV-----ETHNCTKKONRYCACNADSYCALKLHSGNCRQCMKLSKG 126
QY 150 PCTEAEKLDKGVGNHCVCKAGHFONTSSPSARCOPHTRCENOGLVFAAPGTAQSDTT 209
Db 127 PGF-GVARSRTSGNVICSAACAGTFSDDTSTDCVCRPHRCS-----ILAIPGNASTDAV 181
QY 210 CKNPLEPLPEMSSGSLKRRPQGGPNVAGSWEP- PKAHYPFP-----DLVQPLLPIS 262
Db 182 CASE-SPTSAVPRTIYVSQPEPTRSQPM--DQEPGSPQTHIPVSLGSTPIEP--SIT 236
QY 263 GDVS-PVS--TGLPAAPVLEAG-----VPOQSQPLDLTRE---POLEPGSQSVAHGT 309
Db 237 GGISLPIGLIVGLTGLLMLGLANCFLVQRKKKPSCLQRETMPHL--PDEKSDQDAVL 295
QY 310 NGIHV--TGGSMITGNIYINGVPLGGPPGPDLPATPEPPYPIPE 354
Db 296 EQOHLTTAPSSSSSS-----LESSASAGDRRA---PPGVIPQ 330

```

RESULT 2

```

Q62327
ID Q62327 PRELIMINARY; PRT; 459 AA.
AC Q62327;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Murine tumour necrosis factor receptor 2 protein (Fragment).
GN TNFRSF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Amino acid variation in the tumor necrosis factor receptor 2 is
RT linked to autoimmune diabetes in NOD mice."
RL Genomics 0:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=95178848; PubMed=7873884;
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
RT gene."
RL Mamm. Genome 5:726-727(1994).
DR EMBL; X76401; CAA53981.1; -.
DR HSSP; P19438; INCF.
DR MGD; MGI:1314883; Tnfrsf1b.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 4.
DR SMART; SM00208; TNFR; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
KW Receptor.
FT NON_TER 1 1
FT VARIANT 87 87 S -> T.
FT VARIANT 93 93 T -> I.
FT VARIANT 268 268 F -> I.
FT VARIANT 345 345 S -> F.
FT VARIANT 421 421 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CFC14626DF CRC64;

```

Query Match 13.9%; Score 312; DB 11; Length 459;

Best Local Similarity 29.28%; Pred. No. 5.5e-17;

Matches 89; Conservative 41; Mismatches 135; Indels 40; Gaps 13;

```

QY 35 PYASE-NQTCRQDEKEYEYPOHRRICSCRCPPGTYYSAKCSRIKRDIVCATCAENSNEHWN 93
Db 16 PYKPEPGYECQISQ-EYDRAQMCCKACPPGQYVYKHCNKTSIDIVCADCEASMTQWN 74

```

```

QY 94 YLTICQLCR---PCDPVMGLLEEIAPICTSKRKTQCRQCFMCA--AWALECTHCELLSDC 148
Db 75 QFRTCLSCSSCSTDQV---ETRACTKQONRVCAACEAGRYCALKTHSGSCRCQMRLSKC 130
QY 149 PGTEAEKLDKGVGNHCVCKAGHFONTSSPSARCOPHTRCENOGLVFAAPGTAQSDT 208
Db 131 GPGF-GVASSRAPNGNVLCACAPGTFTSDTSTDCVCRPHRCS-----ILAIPGNASTDA 185
QY 209 TCKNPLEPLPEMSSGSLKRRPQGGPNVAGSWEP- PKAHYPFPDVLVQPL----- 258
Db 196 VCA-PESTLTAIPRTLYVSOPEPTRSQPL--DQEPGFSQT--PSILISLSTPIIEOST 240
QY 259 -----LPISGDVSPYSTGLPAAPVLEAG--VPOQSQPLDLTREPOLE--PGSQSVAHGT 309
Db 241 KGGISLPIGLIVGLTGLLMLGLVNCFLVQRKKKPSCLQRDQAKVPHVPDEKSDQDAVL 300
QY 310 NGIHV 314
Db 301 EQOHL 305

```

RESULT 3

```

O88734
ID O88734 PRELIMINARY; PRT; 482 AA.
AC O88734;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE P80 TNF-alpha receptor.
GN TNFR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RX MEDLINE=98414512; PubMed=9740674;
RA Hurle B., Segade F., Rodriguez R., Ramos S.S., Lazo P.S.;
RT "The Mouse Tumor Necrosis Factor Receptor 2 Gene: Genomic Structure
RT and Characterization of the two Transcripts."
RL Genomics 52:79-98(1998).
DR EMBL; Y14619; CAA74969.1; -.
DR EMBL; Y14620; CAA74969.1; JOINED.
DR EMBL; Y14621; CAA74969.1; JOINED.
DR EMBL; Y14622; CAA74969.1; JOINED.
DR EMBL; Y14623; CAA74969.1; JOINED.
DR EMBL; Y14679; CAA74969.1; JOINED.
DR HSSP; P19438; INCF.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 4.
DR SMART; SM00208; TNFR; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
KW Receptor.
SQ SEQUENCE 482 AA; 51106 MW; F6C15046B48FF83C CRC64;

```

Query Match

Best Local Similarity 13.9%; Score 311.5; DB 11; Length 482;

Matches 102; Conservative 48; Mismatches 162; Indels 59; Gaps 17;

```

QY 9 APGLAWPLVLGLFGLLAASQFQAVP-----PYASE-NQTCRQDEKEYEYPOHRRICSR 61
Db 2 APAALWVALV---FELQLWATHTVPAQVLTVPKPEGYECQISQ-EYDRAQMCCKAC 57
QY 62 CPPGTYYSAKCSRI-----RDTVCATCAENSNEHWNLTICOLCR---PCDPVMGLE 111
Db 58 CPPGQYVYKHCNKTSIDIVCADCEASMTQWNQNFRTCLSCSSCSTQDV----- 113
QY 112 EIAPCTSKRKTQCRQCFMCA--AWALECTHCELLSDCPPGTEAEKLDKGVGNHCVCP 169
Db 114 ETTRACTKQONRVCAACEAGRYCALKTHSGSCRCQMRLSKCQPGF-GVASSRAPNGNVLC 172
QY 170 CKAGHFONTSSPSARCOPHTRCENOGLVFAAPGTAQSDTTCNKPLEPLPEMSSGSLKRR 229

```


RESULT 5	
Q8WMQ2	
ID	Q8WMQ2 PRELIMINARY; PRT; 277 AA.
AC	Q8WMQ2;
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Membrane protein CD40 (Fragment).
OS	Ovis aries (Sheep).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Bovidae; Caprinae; Ovis.
OX	NCBI_TaxID=9940;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Zawitkowski M.S., Russ G.R., Krishnan R.;

RESULT	6
Q98SM6	
ID	Q98SM6 PRELIMINARY; PRT; 651 AA.
AC	C98SM6;
DT	01-JUN-2001 (TReMBLrel..17, Created)
DT	01-OCT-2001 (TReMBLrel..18, Last sequence update)
DT	01-DEC-2001 (TReMBLrel..19, Last annotation update)
DE	Death receptor 6.
OS	Gallus gallus (Chicken).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC	Gallus.
NCBI_TaxID=9031;	
[1]	
RN	SEQUENCE FROM N.A.
RP	Bridgham J.T., Johnson A.L.;
RA	"Expression of DR6 in the ovary.";
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF349908; AAK29666.2; -.
DR	HSSP; P19438; 1NCF.
DR	InterPro; IPR000488; Death.
DR	InterPro; IPR003975; Shal_channel.
DR	InterPro; IPR001368; TNFR_c6.
DR	Pfam; PF00531; death; 1.
DR	Pfam; PF00020; TNFR_c6; 4.
DR	PRINTS; PR01497; SHALCHANNEL.
DR	SMART; SM00005; DEATH; 1.
DR	SMART; SM00208; TNFR; 4.
DR	PROSITE; PS00017; DEATH_DOMAIN; 1.
DR	PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR	PROSITE; PS50050; TNFR_NGFR_2; 1.
KW	Receptor.
SQ	SEQUENCE. 651 AA; 71003 MW; BDC95A600DAB2C2A CRC64;
Query Match	11.8%; Score 266; DB 13; Length 651;

Best Local Similarity 25.1%; Pred. No. 3.9e-13;
Matches 107; Conservative 46; Mismatches 150; Indels 124; Gaps 25;

QY 18 VLGLFGLLAASQPAVPPYASENOTCRDQKEYEY-----PQHRRICSRCPGPGTVVSAKCS 73
Db 9 VLPLLVGLTADQ--PKLTSEQANVSLPAGKYLHLDRATNQLICDKCPAGTVVSKHT 66

QY 74 RIRDTVCATCAENSYNHWNLYTICQLCR-PCDPVMGLEEIAPTCKSKRKTQCRQCPGMF- 131
Db 67 KSTLRECSPCPDGFTKHENGIERCHPCRPCE--LPWIEKTHCTALTDRRECTLSGTFF 124

QY 132 ----CAWALECTHCELLSDCP-----PGTEAEKDEVGKGNHCVPCKAGHFQNTSS 180
Db 125 INDTVCVPTV-----CPVGWYRKKGTETE-----DVRCKPCLRGTFSDVPS 166

QY 181 PSARCQPHRCENOGLEAAGTAQSDTTCKNPLEPPMSGLLKRKRRPQGEPPNPVAG 240
Db 167 SVMCKTYTDCFGKMMVVRPGTKESDNCVSPAS-LP---NLSLTSSDAQADETYEA- 221

QY 241 SWEPKHAHYFP-DLVQPLIPISGDVSP-VSTGLPAAVLEAGVPQQOQSLDLTREPQLE 298
Db 222 ---PPTA--YLPKGLNSVFDLSSSPAPRVSN-----TAEPD 256

QY 299 PGESQVAGHTNGIHVTGSGMTITGNI-----YIYNGPVLGPPGP-----GDLPATPE 347
Db 257 YNDTS---ANGTVG---APGSLSSAGTAGQAQSYRKHKTSQAMGKQPAQEMAGGEKSSIPY 311

QY 348 PPYPPIEGDPPGLSTPHQEDGKAWHLAETH-----CCATPSNR 389
Db 312 RP---GPNV---HQ-----HFDINEHLPWMLVFLLLVLIWVVCVSRSSR 356

QY 390 ----GPR 392
Db 357 TLKKGPR 363

RESULT 7

ID Q9DDD2 PRELIMINARY; PRT; 276 AA.
AC Q9DDD2;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Human CD40-homologue.
GN TNFSF5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Tregaskes C.A.;
RL Thesis (2001), University of Reading, Reading, UNITED KINGDOM.
DR EMBL; A293700; CAC2018.1; .
DR HSP; P25942; ICDF.
DR InterPro; IPR001005; Myb_DNA_binding.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 3.
DR SMART; SM00208; TNFR; 3.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
SQ SEQUENCE 276 AA; 30009 MW; 4040B7E0DB82454E CRC64;

Query Match 11.3%; Score 253; DB 13; Length 276;
Best Local Similarity 26.3%; Pred. No. 1.6e-12;
Matches 84; Conservative 31; Mismatches 122; Indels 82; Gaps 13;

QY 19 LGFLGLLAA-----SOPAVPPYASENOTCRDQKEYEYEPQHRICSRCPGTVVSAKCS 73
Db 4 LGLLGLLALLGGQP-----GDAVNCSDQ---YEHKGR-CCNRCQPGKKLASECN 52

QY 74 RIRDTVCATCAENSYNHWNLYTICQLCRPCDPVMGLEEIAPTCKSKRKTQCRQCPGMFCA 133
Db 53 DTDSVCTPCENGQYQHSWTKERHCTPHEICEDNAGLIVRRHGNATHNTVCQCRAGMHCS 112

QY 134 AWALECTHCELLSDCPGP-----TEAEKDEVGKGNHCVPCKAGHFQNTSSPSARCO 186
Db 113 --DASCOTCVENEPCKGFGFVAAAEARM-----TSPCEPCAEGTFSNVSSKTEPCH 163

QY 187 PHTRCENOGLEAAGTAQSDTTCKNP-----LELPPEMGSGL-----LK 227
Db 164 FWTSCEBGLVVKVGINTSDVICESSRRSSLSVLIPITAANVTCLVIGICICLYCLVHTDLR 223

QY 228 RRQGEPPNPVAGSWEPKHAHYFPDLVQPLIPISGDVSPVSTGLPAAVLEAGVPQQOS 287
Db 224 RR-----GPKQAEAEAPR-----ELVTQ-----QPEEVDFPVQETLLGGQVPAQ- 262

QY 288 PLDLTREPQLEPGEOSOVA 306
Db 263 -----EDGESRIA 271

RESULT 8

ID O02764 PRELIMINARY; PRT; 267 AA.
AC O02764;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE OX40 precursor (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHB:HM;
RA Isono T., Seto A.;
RT "Expression of OX40 and OX40 ligand genes in rabbit HTLV-I-transformed T cell lines."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003911; BRA20059.1; .
DR HSP; P25942; ICDF.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 3.
DR SMART; SM00208; TNFR; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 18 POTENTIAL.
FT CHAIN 19 267 OX40.
SQ SEQUENCE 267 AA; 28489 MW; A8B4CD3173C9500B CRC64;

Query Match 11.2%; Score 252.5; DB 6; Length 267;
Best Local Similarity 26.7%; Pred. No. 1.7e-12;
Matches 75; Conservative 32; Mismatches 95; Indels 79; Gaps 11;

QY 17 LVGLFGLLAAASQPAVPPYASENOTCRDQKEYEYEPQHRICSRCPGTVVSAKCSRIR 76
Db 8 LALLLGLLIGAEPR--PDCVGDTPGDR-----CCLECPQGYGMYSCRNSQ 54

QY 77 DTVCATCAENSYNHWNLYTICQLCRPCDPVMGLEEIAPTCKSKRKTQCRQCPGMFCAAWA 136
Db 55 DTICHPCPGFYNEAVNY-QACKPCTCNRRSGSEPOQECTHTRTDTCRCRPG----- 106

QY 137 LECTHCELLSDCPGTEAEKDEVGKGNHCVPCKAGHFQNTSSPSARCPHTCENQGL 196
Db 107 -----TQPLNGYKHGVD-----CAPCQGHFSEGNRA--CRPTNCTLAGK 146

QY 197 VEAPGTAQSDTTCKNPLEPPMSGLLKRKRRPQGEPPNPVAGSWEPKHAHYFPDLVQ 256
Db 147 RTLOPASSISDANCED-----RSSLATQP-----WETPSA-PYRPTAR 184


```
Db 70 TKERCPPHTYTPVNSYSGCHQCRKC-PTGSFDKV-KCTGTQNSKCSCLPGWFCATDSS 127
QY 135 WALECTHCELLSDCPGTEAEKDEVGKGNHCVCKKAGHF 175
Db 128 KYEDCRDCLPKRCKPCGCGYGGI-DEL--GNPLCKSCCVGEY 165

RESULT 14
O57117 PRELIMINARY; PRT; 351 AA.
AC O57117;
DT 01-JUN-1998 (TremBLrel. 06, Created)
DT 01-JUN-1998 (TremBLrel. 06, Last sequence update)
DE 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Tumor necrosis factor receptor II homolog.
GN CRMB.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MUNICH OPV 85 (HUMAN);
RX MEDLINE=20579014; PubMed=11136755;
RA Loparev V.N., Massung R.F., Esposito J.J., Meyer H.;
RT "Detection and differentiation of old world orthopoxviruses:
RL restriction fragment length polymorphism of the crmb gene region.";
RL J. Clin. Microbiol. 39:94-100(2001).
DR EMBL; U90227; A894383.1; -
DR HSP; O14763; IDOG.
DR InterPro; IPR001368; TNFR_c6.
DR Pfam; PF00020; TNFR_c6; 2.
DR SMART; SM00208; TNFR_2.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
KW Receptor.
SQ SEQUENCE 351 AA; 38304 MW; 57C84FFE5B5F57E CRC64;

Query Match 10.2%; Score 230; DB 12; Length 351;
Best Local Similarity 31.4%; Pred. No. 1.5e-10;
Matches 50; Conservative 18; Mismatches 79; Indels 12; Gaps 6;

QY 33 VPPYASENCTDQKEYYEPQHRRICSCRCPPGTYVSARK--SRIRDTVCATCAENSYNE 90
Db 22 VAPYAPSNCKRCDNEYN----RHNLCLSCPPGTYASRLCDSKTYNTQCTPGSGCTFTS 77
QY 91 HNNYLTICOLR-PCDPVNGLEEIAPTCKRKTQCRQPMFC-AAWALECTHCELLSDC 148
Db 78 HNNHLPACLSNCRCD--SNQVETPCNTTHNRICSPGYCYLLKGGSGCKACVSTKRC 135
QY 149 PCTEAEKDEVGKGNHCVCKKAGHFQNTSSPSARCP 187
Db 136 --GIGYVSGHTSTGDVICSPCGLGTYSHTVSSADKCEP 172

RESULT 15
Q90YS6 PRELIMINARY; PRT; 285 AA.
AC Q90YS6;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE TNF decoy receptor.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu L., Fujiki K., Dixon B., Sundick R.S.;
RT "Cloning of a novel rainbow trout (Oncorhynchus mykiss) CC chemokine
```

```
RT with a fractalkine-like stalk and a TNF decoy receptor using cdna
RT fragments containing AU-rich elements.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF401631; AAK91758.1; -
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001368; TNFR_c6; 3.
DR Pfam; PF00020; TNFR_c6; 3.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 285 AA; 31795 MW; 5E3BD1B6EFC6BABC CRC64;

Query Match 10.1%; Score 228; DB 13; Length 285;
Best Local Similarity 29.6%; Pred. No. 1.6e-10;
Matches 61; Conservative 18; Mismatches 87; Indels 40; Gaps 10;

QY 17 LVILGLFGLLAASQQA---VPPYASENCTDQKEYYEPQHRRICSCRCPPGTYVSARKS 73
Db 9 LPLLVFALCGSGVPVSGAHTPTTYWRDDATGDS-----LTCDLCAPGTYLLKHCT 58
QY 74 RIRDTVCATCAENSYNEHNNYLTICOLC-RPC--DPVNGLEEIAPTCKRKTQCRQPM 130
Db 59 KDRKSDCGPCPKSHYTEIWNYYIERQYCNRFCTADEI----ESVPCTQLHNRQCECKDGF 114
QY 131 FCAAWALECTH--CELLSDCPGTEAEKDEVGKGNH---CVPCKAGHFQNTSSPSAR 184
Db 115 Y-----MTHGSCSRHRCPPG-----EGVISNGTAHTDVKCEPCVGVFFSAVSSRKA 162
QY 185 CQPHTRCENOGLEAAPGTAQSDYTC 210
Db 163 CQKFSVCPPGR--TTIPGNDMDVYC 186

Search completed: April 8, 2003, 10:51:35
Job time : 105.845 secs
```


Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1404	62.5	415	4	US-09-006-353A-6	Sequence 6, Appli
2	1404	62.5	415	4	US-09-573-986-6	Sequence 6, Appli
3	1129	50.2	197	2	US-08-505-608-1	Sequence 1, Appli
4	1129	50.2	197	4	US-09-000-166-1	Sequence 1, Appli
5	987	43.9	170	4	US-08-828-683A-14	Sequence 14, Appl
6	456	20.3	77	4	US-08-866-54-3	Sequence 3, Appli
7	349.5	15.5	461	1	US-08-385-229-2	Sequence 2, Appli
8	349.5	15.5	461	2	US-08-650-000-2	Sequence 2, Appli
9	349.5	15.5	461	4	US-09-042-785A-7	Sequence 7, Appli
10	349.5	15.5	461	4	US-08-477-347-3	Sequence 3, Appli
11	349.5	15.5	461	4	US-09-006-353A-4	Sequence 4, Appli
12	349.5	15.5	461	4	US-08-476-862-2	Sequence 2, Appli
13	349.5	15.5	461	4	US-09-573-986-4	Sequence 4, Appli
14	349.5	15.5	461	6	5395760-2	Patent No. 5395760
15	339.5	15.1	518	1	US-08-385-229-4	Sequence 4, Appli
16	337.5	15.0	486	1	US-08-243-010-1	Sequence 1, Appli
17	327.5	14.6	474	2	US-08-650-000-4	Sequence 4, Appli
18	327.5	14.6	474	4	US-09-042-785A-8	Sequence 8, Appli
19	327.5	14.6	474	6	5395760-4	Patent No. 5395760
20	321.5	14.3	227	3	US-08-974-022-48	Sequence 48, Appli
21	321.5	14.3	227	4	US-08-795-445A-48	Sequence 48, Appli
22	321.5	14.3	227	4	US-08-795-447A-48	Sequence 48, Appli
23	321.5	14.3	227	4	US-08-974-186-48	Sequence 48, Appli
24	321.5	14.3	227	4	US-08-795-446B-48	Sequence 48, Appli
25	321.5	14.3	227	4	US-08-706-945D-134	Sequence 134, Appli
26	319	14.2	300	2	US-08-794-796-2	Sequence 2, Appli
27	314.5	14.0	276	4	US-09-041-886-27	Sequence 27, Appli

Db 297 PAGPTAPSLVEVLQOQSPVQARELEAPGEGHGVAGHANGIHVTGGSVTVTGNIIY 356

QY 329 NGVILGPPGCDLPATPEPPYPIPEGDPGPPGLSTPHOEDGKAWHLAETEHCG 383

Db 357 NGPVLGTRGPGDPAPPPEPYTPPEGAPGPPSELSTPYQEDGKAWHLAETETLG 411

RESULT 3

US-08-505-606-1

; Sequence 1, Application US/08505606

; Patent No. 5925351

; GENERAL INFORMATION:

; APPLICANT: BROWNING, Jeffrey L.

; APPLICANT: BROWNING, Christopher D.

; APPLICANT: HOCHEMAN, Paula S.

; TITLE OF INVENTION: SOLUBLE LYMPHOTOXIN-BETA RECEPTORS AND

; TITLE OF INVENTION: ANTI-LYMPHOTOXIN RECEPTOR AND LIGAND ANTIBODIES AS

; TITLE OF INVENTION: THERAPEUTIC AGENTS FOR THE TREATMENT OF IMMUNOLOGICAL

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: James F. Haley, Jr.

; STREET: 1251 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10020

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/505,606

; FILING DATE: 21-JUL-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/378,968

; FILING DATE: 26-JAN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: HALEY, Jr., James F.

; REGISTRATION NUMBER: 27,794

; REFERENCE/DOCKET NUMBER: B191

; TELEPHONE: (212) 596-9000

; TELEFAX: (212) 596-9090

; TELEX: 14-8367

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 197 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-505-606-1

Query Match 50.2%; Score 1129; DB 2; Length 197;

Best Local Similarity 99.5%; Pred. No. 2.3e-75;

Matches 196; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 28 SQQAVPVYASNTQCRDOEKEYEYEPQHRICCSRCPPTGYYSAKCSRIKRDIVCATCAENS 87

Db 1 SQQAVPVYASNTQCRDOEKEYEYEPQHRICCSRCPPTGYYSAKCSRIKRDIVCATCAENS 60

QY 88 YNEHWNLTICQLRCPDPMVGLIEIAPCTSKRKTQCRCPGMFCAAWALECTHCELLSD 147

Db 61 YNEHWNLTICQLRCPDPMVGLIEIAPCTSKRKTQCRCPGMFCAAWALECTHCELLSD 120

QY 148 CPPGTAEALKDEVGKNNHCVPCKAGHFQNTSSPACQPHTRCENQGLVEAPGTAQSD 207

Db 121 CPPGTAEALKDEVGKNNHCVPCKAGHFQNTSSPACQPHTRCENQGLVEAPGTAQSD 180

QY 208 TTCKNPLEPLPPMMSG 224

QY 61 RCPPTGYYSAKCSRIKRDIVCATCAENS YNEHWNLTICQLRCPDPMVGLIEIAPCTSKR 120

Db 61 RCPPTGYYSAKCSRIKRDIVCATCAENS YNEHWNLTICQLRCPDPMVGLIEIAPCTSKR 120

QY 121 KTCRCQPGMFCFAAWALECTHC--ELLSDCPGTEAEALKDEVGKNNHCVPCKAGHFQNT 178

Db 121 KTCRCQPGMFCFAAWALECTHC--ELLSDCPGTEAEALKDEVGKNNHCVPCKAGHFQNT 180

QY 179 SSPARCOPTHCENQGLVEAPGTAQSDTTCKNPLEP----- 216

Db 181 SSPARCOPTHCENQGLVEAPGTAQSDTTCKNPLEP----- 240

QY 217 -----LPPEMS---GSLKRRPQEGPNPVAGSWEPKKAHYFPDLPVLPISGDSVPV 268

Db 241 ACAMWRHPSLCRKLGTLLKRHPGEESPPCPA---PRADPHFPDLAEPPLPMSGDLSPS 296

QY 269 STGLPAAPVLEAGVPOOQSLDITREPQLEGEQSQVAHGTHGIVTGGSMITGNIIY 328

Db 297 PAGPTAPSLVEVLQOQSPVQARELEAPGEGHGVAGHANGIHVTGGSVTVTGNIIY 356

QY 329 NGVILGPPGCDLPATPEPPYPIPEGDPGPPGLSTPHOEDGKAWHLAETEHCG 383

Db 357 NGPVLGTRGPGDPAPPPEPYTPPEGAPGPPSELSTPYQEDGKAWHLAETETLG 411

RESULT 2

US-09-573-986-6

; Sequence 6, Application US/09573986

; Patent No. 6455040

; GENERAL INFORMATION:

; APPLICANT: Wei, Ying-Fei

; APPLICANT: Ni, Jian

; APPLICANT: Gentz, Reiner

; APPLICANT: Ruben, Steven

; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5

; FILE REFERENCE: 1488.1280004

; CURRENT APPLICATION NUMBER: US/09/573,986

; CURRENT FILING DATE: 2000-05-18

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 6

; LENGTH: 415

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-573-986-6

Query Match 62.5%; Score 1404; DB 4; Length 415;

Best Local Similarity 63.9%; Pred. No. 4.7e-95;

Matches 265; Conservative 28; Mismatches 86; Indels 36; Gaps 4;

QY 1 MLPLWATAPGLAWGLPLVLGLFLLAASQAVPVYASNTQCRDOEKEYEYEPQHRICCS 60

Db 1 MLPLWATAPGLAWGLPLVLGLFLLAASQAVPVYASNTQCRDOEKEYEYEPQHRICCS 60

QY 61 RCPPTGYYSAKCSRIKRDIVCATCAENS YNEHWNLTICQLRCPDPMVGLIEIAPCTSKR 120

Db 61 RCPPTGYYSAKCSRIKRDIVCATCAENS YNEHWNLTICQLRCPDPMVGLIEIAPCTSKR 120

QY 121 KTCRCQPGMFCFAAWALECTHC--ELLSDCPGTEAEALKDEVGKNNHCVPCKAGHFQNT 178

Db 121 KTCRCQPGMFCFAAWALECTHC--ELLSDCPGTEAEALKDEVGKNNHCVPCKAGHFQNT 180

QY 179 SSPARCOPTHCENQGLVEAPGTAQSDTTCKNPLEP----- 216

Db 181 SSPARCOPTHCENQGLVEAPGTAQSDTTCKNPLEP----- 240

QY 217 -----LPPEMS---GSLKRRPQEGPNPVAGSWEPKKAHYFPDLPVLPISGDSVPV 268

Db 241 ACAMWRHPSLCRKLGTLLKRHPGEESPPCPA---PRADPHFPDLAEPPLPMSGDLSPS 296

QY 269 STGLPAAPVLEAGVPOOQSLDITREPQLEGEQSQVAHGTHGIVTGGSMITGNIIY 328

|||||
Db 181 TTCKNPLELPPEMSGT 197
RESULT 4
US-09-000-166-1
; Sequence 1, Application US/09000166A
; Patent No. 6403087
; GENERAL INFORMATION:
; APPLICANT: Browning, et al.
; TITLE OF INVENTION: Soluble Lymphotoxin-B Receptors and Anti-lymphotoxin
; TITLE OF INVENTION: Receptor and Ligand Antibodies, as Therapeutic Agents
; TITLE OF INVENTION: for the treatment of Immunological Disease.
; FILE REFERENCE: B191
; CURRENT APPLICATION NUMBER: US/09/000,166A
; EARLIER FILING DATE: 1998-06-08
; EARLIER APPLICATION NUMBER: PCT/US96/12010
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-000-166-1
Query Match 50.2%; Score 1129; DB 4; Length 197;
Best Local Similarity 99.5%; Pred. No. 2.3e-75;
Matches 196; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 28 SQQAVPPYASENOTCRDQKEYEYEPQHRICCSRCPPGTVVSAAKSRIRDTVCATCAENS 87
Db 1 SQQAVPPYASENOTCRDQKEYEYEPQHRICCSRCPPGTVVSAAKSRIRDTVCATCAENS 60
QY 88 YNEHNYLTICQLCRPCDPVGMLEETAPCTSKRKTQCRQCPGMFCAAWALECTHCELLSD 147
Db 61 YNEHNYLTICQLCRPCDPVGMLEETAPCTSKRKTQCRQCPGMFCAAWALECTHCELLSD 120
QY 148 CPPGTEAEKLDKGVGNHCVCKAGHFONTSSPSARCQPHTRCENOGLVAAAPGTAQSD 207
Db 121 CPPGTEAEKLDKGVGNHCVCKAGHFONTSSPSARCQPHTRCENOGLVAAAPGTAQSD 180
QY 208 TTCKNPLELPPEMSGS 224
Db 181 TTCKNPLELPPEMSGT 197
RESULT 5
US-08-828-683A-14
; Sequence 14, Application US/08828683A
; Patent No. 6469144
; GENERAL INFORMATION:
; APPLICANT: Askenazi, Avi J.
; TITLE OF INVENTION: Apo-2 LI AND Apo-3 POLYPEPTIDES
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/828,683A
; FILING DATE: 31-Mar-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/625328

;; FILING DATE: 1-Apr-1996
;; APPLICATION NUMBER: 08/710802
;; FILING DATE: 23-Sep-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Marschang, Diane L.
;; REGISTRATION NUMBER: 35,600
;; REFERENCE/DOCKET NUMBER: P1007P1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650/225-5416
;; TELEFAX: 650/952-9881
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 170 amino acids
;; TYPE: Amino Acid
;; TOPOLOGY: Linear
;; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-08-828-683A-14
Query Match 43.9%; Score 987; DB 4; Length 170;
Best Local Similarity 100.0%; Pred. No. 4.4e-65;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 TCRDQKEYEYEPQHRICCSRCPPGTVVSAAKSRIRDTVCATCAENSNEHNYLTICQLC 101
Db 1 TCRDQKEYEYEPQHRICCSRCPPGTVVSAAKSRIRDTVCATCAENSNEHNYLTICQLC 60
QY 102 RPCDPVGMLEETAPCTSKRKTQCRQCPGMFCAAWALECTHCELLSDCPPGTEAEKLDKVG 161
Db 61 RPCDPVGMLEETAPCTSKRKTQCRQCPGMFCAAWALECTHCELLSDCPPGTEAEKLDKVG 120
QY 162 KGNHCVCKAGHFONTSSPSARCQPHTRCENOGLVAAAPGTAQSDTTCK 211
Db 121 KGNHCVCKAGHFONTSSPSARCQPHTRCENOGLVAAAPGTAQSDTTCK 170
RESULT 6
US-08-866-545-3
; Sequence 3, Application US/08866545
; Patent No. 6265535
; GENERAL INFORMATION:
; APPLICANT: Greene, Mark I.
; APPLICANT: Murall, Ramachandran
; APPLICANT: Takasaki, Wataru
; TITLE OF INVENTION: PEPTIDES AND PEPTIDE
; TITLE OF INVENTION: ANALOGUES DESIGNED FROM BINDING SITES OF TUMOR
; TITLE OF INVENTION: NECROSIS FACTOR RECEPTOR SUPERFAMILY AND THEIR
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,545
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 009113-0004-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935

; REFERENCE/DOCKET NUMBER: 2501-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-650-000-2

Query Match 15.5%; Score 349.5; DB 2; Length 461;
Best Local Similarity 29.7%; Pred. No. 4.3e-18;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APGLAWPLVLGLFLLAASQPAVP-----PYASE-NQCRDQEKYEYEPQHRICCSR 61
DB 2 APVAVMAALAVGL-ELWAAA--HALPAQVAFTPYAPEPGSTCR--LREYDQTAQMCCSK 56
QY 62 CPPGTYSVSAKCSRIKDTVCATCAENSYNEHNNYLITCOLCR---PCDPVMGLEETAPCTS 118
DB 57 CSPGHAKVFCIKTSDTVCDSCEDSTYTLNWNWPECLSCGSRCSDDV---ETOACTR 112
QY 119 KRKTQCRQCPGMFCAAWALE-CTHCELLSDCPPG-----TEAELKDEYKGNHNCVPCA 172
DB 113 EQNRICTRCPGWYCALSKQEGCRCLCAPLRCRPGFVGVARPGTETSDVV-----CKPCAP 166
QY 173 GHFQNTSSPSARCQPHTRCENGLVEAAPGTAQSDTTC--KNPLEPLPPMSGSL----- 225
DB 167 GTFSNTTSTDICRPHQICN-----VVAIPGNASMDAVCTSTSPTRSMAP---GAYHLPQP 219
QY 226 LKRRPQEGPNVAGSWPEPPKAH-----PYFPDLVQPLLPISGDIVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPT-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVLGLVGTALG 272
QY 277 VLEAGVQP-----QQSPDLTREPQLE--PGEQSOVAHGNTGTHV-----TGGNMT 320
DB 273 LLIIIGVNCVIMTVQKKKPLCLQREAKVPHLPADKARGTQGPQQHLLITAPSSSSSLE 332
QY 321 ITGNIYIYNGPVILGPPGPG 340
DB 333 SSASALDRAPTRNQPPAG 352

RESULT 9
US-09-042-785A-7
; Sequence 7, Application US/09042785A
; Patent No. 6194151

; GENERAL INFORMATION:
; APPLICANT: Busfield, Samantha J
; TITLE OF INVENTION: NOVEL MOLECULES OF THE TNF RECEPTOR SUPERFAMILY
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; APPLICATION NUMBER: US/09/042.785A
; FILING DATE: 17-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/938,896
; FILING DATE: 26-SEP-1997
; ATTORNEY/AGENT INFORMATION:

; NAME: Mandragouras, Amy E
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: MEI-001CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-042-785A-7

Query Match 15.5%; Score 349.5; DB 4; Length 461;
Best Local Similarity 29.7%; Pred. No. 4.3e-18;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APGLAWPLVLGLFLLAASQPAVP-----PYASE-NQCRDQEKYEYEPQHRICCSR 61
DB 2 APVAVMAALAVGL-ELWAAA--HALPAQVAFTPYAPEPGSTCR--LREYDQTAQMCCSK 56
QY 62 CPPGTYSVSAKCSRIKDTVCATCAENSYNEHNNYLITCOLCR---PCDPVMGLEETAPCTS 118
DB 57 CSPGHAKVFCIKTSDTVCDSCEDSTYTLNWNWPECLSCGSRCSDDV---ETOACTR 112
QY 119 KRKTQCRQCPGMFCAAWALE-CTHCELLSDCPPG-----TEAELKDEYKGNHNCVPCA 172
DB 113 EQNRICTRCPGWYCALSKQEGCRCLCAPLRCRPGFVGVARPGTETSDVV-----CKPCAP 166
QY 173 GHFQNTSSPSARCQPHTRCENGLVEAAPGTAQSDTTC--KNPLEPLPPMSGSL----- 225
DB 167 GTFSNTTSTDICRPHQICN-----VVAIPGNASMDAVCTSTSPTRSMAP---GAYHLPQP 219
QY 226 LKRRPQEGPNVAGSWPEPPKAH-----PYFPDLVQPLLPISGDIVS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPT-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVLGLVGTALG 272
QY 277 VLEAGVQP-----QQSPDLTREPQLE--PGEQSOVAHGNTGTHV-----TGGNMT 320
DB 273 LLIIIGVNCVIMTVQKKKPLCLQREAKVPHLPADKARGTQGPQQHLLITAPSSSSSLE 332
QY 321 ITGNIYIYNGPVILGPPGPG 340
DB 333 SSASALDRAPTRNQPPAG 352

RESULT 10
US-08-477-347-3
; Sequence 3, Application US/08477347
; Patent No. 6232446
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: BIGDA, Jacek
; APPLICANT: BELETSKY, Igor
; APPLICANT: METT, Igor
; TITLE OF INVENTION: TNF LIGANDS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,347

; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/115,685
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: IL 106271
 ; FILING DATE: 08-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Townsend, G. Kevin
 ; REGISTRATION NUMBER: 34,033
 ; REFERENCE/DOCKET NUMBER: WALLACH-10
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-628-5197
 ; TELEFAX: 202-737-3528
 ; TELEX: 248633
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 461 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-477-347-3

Query Match 15.5%; Score 349.5; DB 4; Length 461;
 Best Local Similarity 29.7%; Pred. No. 4.3e-18;
 Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APLGAWPLVLGLGLAASQPOAVP-----PYASE-NOTCRDQKEYYEQHRIICCSR 61
 Db 2 APVAVWAALAVGL-ELWAAA--HALPAQVAFTPYAPPGSTCR--LREYDQTAQMCCSK 56
 QY 62 CPPTVYSKACSRIRDTVCATCAENSYNEHWNLTICQLCR---PCDPVMGLEEIAPICTS 118
 Db 57 CSPGQHAKEVCTKSTVTCVDCSDSTYTQLWNVPECLSCGSRCSQDV---ETQACTR 112
 QY 119 KRKTQRCQCPGMFCAAWALE-CTHCELLSDCPGP-----TEALKDEVGKNNHCVPCKA 172
 Db 113 EQNRICTCRPGWCALSKQEGRCRLAPLKRCPGFGVARGTSDV-----CKPCAP 166
 QY 173 GHFQNTSSPACOPHTRCENGLVEAPGTAQSDTTC--KNPLEPLPPMMSGSL----- 225
 Db 167 GTFSTNTSSDTCRPHQICN---VVAIPGNASMDAVCTSTSTRMAP---GAVHLPQP 219
 QY 226 LKRRPQEGPNPVAGSWEPKHAH-----PYFPDLVQPLLPISGDVS-PVS---TGLPAAP 276
 Db 220 VSTRSQTHTPT-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVLIVGVGTALG 272
 QY 277 VLEAGVPQ-----QOSPLDLTREPQLE--PGEQSOVAHGTINGIHV-----TGSMT 320
 Db 273 LLIIGVNCVIMTVQKKKPLCLQREAKVPHLPADKARGTQGPQQHLLITAPSSSSSLE 332
 QY 321 ITGNIYINGPVLGPPGPG 340
 Db 333 SSASALDRRAPTRNQPAQPG 352

RESULT 11
 US-09-006-353A-4
 ; Sequence 4, Application US/09006353A
 ; Patent No. 6261801
 ; GENERAL INFORMATION:
 ; APPLICANT: WEI, YING-FEI
 ; APPLICANT: YU, GUO-LIANG
 ; APPLICANT: GENTZ, REINER
 ; APPLICANT: RUBEN, STEVEN
 ; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR 5
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
 ; STREET: 9410 KEY WEST AVENUE
 ; CITY: ROCKVILLE
 ; STATE: MD

; COUNTRY: US
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/006,353A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BROOKES, ANDERS A
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PF341
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 461 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-006-353A-4

Query Match 15.5%; Score 349.5; DB 4; Length 461;
 Best Local Similarity 29.7%; Pred. No. 4.3e-18;
 Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APLGAWPLVLGLGLAASQPOAVP-----PYASE-NOTCRDQKEYYEQHRIICCSR 61
 Db 2 APVAVWAALAVGL-ELWAAA--HALPAQVAFTPYAPPGSTCR--LREYDQTAQMCCSK 56
 QY 62 CPPTVYSKACSRIRDTVCATCAENSYNEHWNLTICQLCR---PCDPVMGLEEIAPICTS 118
 Db 57 CSPGQHAKEVCTKSTVTCVDCSDSTYTQLWNVPECLSCGSRCSQDV---ETQACTR 112
 QY 119 KRKTQRCQCPGMFCAAWALE-CTHCELLSDCPGP-----TEALKDEVGKNNHCVPCKA 172
 Db 113 EQNRICTCRPGWCALSKQEGRCRLAPLKRCPGFGVARGTSDV-----CKPCAP 166
 QY 173 GHFQNTSSPACOPHTRCENGLVEAPGTAQSDTTC--KNPLEPLPPMMSGSL----- 225
 Db 167 GTFSTNTSSDTCRPHQICN---VVAIPGNASMDAVCTSTSTRMAP---GAVHLPQP 219
 QY 226 LKRRPQEGPNPVAGSWEPKHAH-----PYFPDLVQPLLPISGDVS-PVS---TGLPAAP 276
 Db 220 VSTRSQTHTPT-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVLIVGVGTALG 272
 QY 277 VLEAGVPQ-----QOSPLDLTREPQLE--PGEQSOVAHGTINGIHV-----TGSMT 320
 Db 273 LLIIGVNCVIMTVQKKKPLCLQREAKVPHLPADKARGTQGPQQHLLITAPSSSSSLE 332
 QY 321 ITGNIYINGPVLGPPGPG 340
 Db 333 SSASALDRRAPTRNQPAQPG 352

RESULT 12
 US-08-476-862-2
 ; Sequence 2, Application US/08476862
 ; Patent No. 6262239
 ; GENERAL INFORMATION:
 ; APPLICANT: WALLACH, David
 ; APPLICANT: BIGDA, Jacek
 ; APPLICANT: BELETSKY, Igor
 ; APPLICANT: METT, Igor
 ; APPLICANT: ENGELMANN, Hartmut
 ; TITLE OF INVENTION: TNF INHIBITORS
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,862
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 107267
FILING DATE: 12-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 94039
FILING DATE: 06-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 91229
FILING DATE: 06-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 90339
FILING DATE: 18-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH-12A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 461 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-862-2

Query Match 15.5%; Score 349.5; DB 4; Length 461;
Best Local Similarity 29.7%; Pred. No. 4.3e-18;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APLGAWPLVLGLFLLAASQPAVP-----PYASE-NQTCROQEKEYEPPQHRICCSR 61
DB 2 APVAVWAALAVGL-ELWAAA--HALPAQVAFPTYAPEPGSTCR--LREYDQTAQMCCSK 56
QY 62 CPPGTIVYSAKCSRIRDTVCATCAENSNEHWNLTICQLCR---PCDPVMGLEETAPCTS 118
DB 57 CSPQOHAKVFTKTSDFVDCSDSTYTLWNWPECLSCGSRSSDVV-----ETQACTR 112
QY 119 KRKTQCRQCPGMFCAAWALE-CTHCELLSDCPPG-----TEAEKDEYKGNHCVCPCKA 172
DB 113 EONRICTCRPGWCYKALSKQEGCRLCAPLRCRPGFVGARPGTETSDVV-----CKPCAP 166
QY 173 GHFQNTSSPSARCOPHRCENOGIVEAPGTAQSDTTC--KNPLEPLPPMSGSL----- 225
DB 167 GTFSNTTSSDTCRPHQICN---VVAIPGNASMDAVCTSTSPTRSMAP---GAVHLPQP 219
QY 226 LKRRPQEGGNPNVAGSNPEPKAH-----PYFPDLVQPLLPISGDVPS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVLGVGTALG 272
QY 277 VLEAGVPO-----QSQPLDLTREPOLE--PGEQSQVAHGTNGIHV-----TGSMT 320
DB 273 LLIGVYVNCVIMTQVKKKPLCLQREAKVPHLPADKARGTQGEQOQHLLITAPSSSSSLE 332
QY 321 ITGNIYYNGVPLGGPPGPG 340
DB 333 SSASALDRRAPTRNQPAQP 352

RESULT 13
US-09-573-986-4
Sequence 4, Application US/09573986
Patent No. 6455040
GENERAL INFORMATION:
APPLICANT: Wei, Ying-Fei
APPLICANT: Ni, Jian
APPLICANT: Gentz, Reiner
APPLICANT: Ruben, Steven
TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
FILE REFERENCE: 1488.1280004
CURRENT APPLICATION NUMBER: US/09/573,986
CURRENT FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 461
TYPE: PRT
ORGANISM: Homo sapiens
US-09-573-986-4

Query Match 15.5%; Score 349.5; DB 4; Length 461;
Best Local Similarity 29.7%; Pred. No. 4.3e-18;
Matches 113; Conservative 47; Mismatches 143; Indels 77; Gaps 22;

QY 9 APLGAWPLVLGLFLLAASQPAVP-----PYASE-NQTCROQEKEYEPPQHRICCSR 61
DB 2 APVAVWAALAVGL-ELWAAA--HALPAQVAFPTYAPEPGSTCR--LREYDQTAQMCCSK 56
QY 62 CPPGTIVYSAKCSRIRDTVCATCAENSNEHWNLTICQLCR---PCDPVMGLEETAPCTS 118
DB 57 CSPQOHAKVFTKTSDFVDCSDSTYTLWNWPECLSCGSRSSDVV-----ETQACTR 112
QY 119 KRKTQCRQCPGMFCAAWALE-CTHCELLSDCPPG-----TEAEKDEYKGNHCVCPCKA 172
DB 113 EONRICTCRPGWCYKALSKQEGCRLCAPLRCRPGFVGARPGTETSDVV-----CKPCAP 166
QY 173 GHFQNTSSPSARCOPHRCENOGIVEAPGTAQSDTTC--KNPLEPLPPMSGSL----- 225
DB 167 GTFSNTTSSDTCRPHQICN---VVAIPGNASMDAVCTSTSPTRSMAP---GAVHLPQP 219
QY 226 LKRRPQEGGNPNVAGSNPEPKAH-----PYFPDLVQPLLPISGDVPS-PVS--TGLPAAP 276
DB 220 VSTRSQHTQPTP-----EPSTAPSTSFLLPMGPS--PPAEGSTGDFALPVLGVGTALG 272
QY 277 VLEAGVPO-----QSQPLDLTREPOLE--PGEQSQVAHGTNGIHV-----TGSMT 320
DB 273 LLIGVYVNCVIMTQVKKKPLCLQREAKVPHLPADKARGTQGEQOQHLLITAPSSSSSLE 332
QY 321 ITGNIYYNGVPLGGPPGPG 340
DB 333 SSASALDRRAPTRNQPAQP 352

RESULT 14
5395760-2
Patent No. 5395760
APPLICANT: SMITH, CRAIG A.; GOODWIN, RAYMOND G.; BECKMANN, M. PATRICIA
TITLE OF INVENTION: DNA ENCODING TUMOR NECROSIS FACTOR-a AND B-RECEPTORS
NUMBER OF SEQUENCES: 17
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/523,635
FILING DATE: 10-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 421,417
FILING DATE: 13-OCT-1989
APPLICATION NUMBER: 405,370
FILING DATE: 11-SEP-1989
APPLICATION NUMBER: 403,241

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:45:16 ; Search time 5.15513 seconds
(without alignments)
799.386 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235

Perfect score: 108

Sequence: 1 PLPPEMSGSLKRRPQGGP 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	46.3	125	17 Q9YCW4	Q9YCW4 aeropyrum p
2	49	45.4	139	11 Q9CTN9	Q9CTN9 mus musculus
3	49	45.4	688	5 Q9W111	Q9W111 drosophila
4	49	45.4	891	5 Q9VGJ8	Q9VGJ8 drosophila
5	49	45.4	1253	5 Q8WQ61	Q8WQ61 drosophila
6	49	45.4	2157	4 Q95875	Q95875 homo sapien
7	49	45.4	2157	4 Q960C6	Q960C6 homo sapien
8	48	44.4	930	11 Q925G9	Q925G9 mus musculus
9	47.5	44.0	1200	10 Q9AR25	Q9AR25 linum usita
10	47	43.5	618	10 Q9MA81	Q9MA81 arabidopsis
11	47	43.5	820	10 Q8RXZ0	Q8RXZ0 arabidopsis
12	47	43.5	1055	10 Q9MA51	Q9MA51 arabidopsis
13	47	43.5	1542	5 Q9VE07	Q9VE07 drosophila
14	46.5	43.1	511	6 Q95J00	Q95J00 sus scrofa
15	46.5	43.1	566	6 Q95J01	Q95J01 sus scrofa
16	46.5	43.1	572	16 Q33009	Q33009 mycobacteri

17	46.5	43.1	676	6 Q95JC9	Q95JC9 sus scrofa
18	46	42.6	471	10 Q49326	Q49326 arabidopsis
19	46	42.6	501	16 Q55762	Q55762 synechocyst
20	46	42.6	512	16 Q8XQ09	Q8XQ09 raietonia s
21	46	42.6	528	10 Q92PL6	Q92PL6 nicotiana t
22	46	42.6	598	4 Q8TBS5	Q8TBS5 homo sapien
23	46	42.6	645	10 Q94LV8	Q94LV8 oryza sativ
24	46	42.6	667	5 Q9Y174	Q9Y174 trichomonas
25	46	42.6	820	4 Q00236	Q00236 homo sapien
26	46	42.6	965	10 Q8S2E8	Q8S2E8 oryza sativ
27	46	42.6	1124	4 Q9Y6S2	Q9Y6S2 homo sapien
28	46	42.6	1124	4 Q99699	Q99699 homo sapien
29	46	42.6	1543	11 Q8R4H2	Q8R4H2 mus musculu
30	46	42.6	1544	4 Q9NZN5	Q9NZN5 homo sapien
31	45.5	42.1	207	13 Q9NL04	Q9NL04 eptatretus
32	45.5	42.1	363	11 Q62314	Q62314 mus musculu
33	45.5	42.1	754	11 Q9JJJ4	Q9JJJ4 mus musculu
34	45.5	42.1	754	11 Q99MU9	Q99MU9 mus musculu
35	45.5	42.1	754	11 Q9Z0R0	Q9Z0R0 mus musculu
36	45.5	42.1	1042	5 Q9V8C2	Q9V8C2 drosophila
37	45.5	42.1	1043	5 Q8T8P9	Q8T8P9 drosophila
38	45.5	42.1	1963	6 Q28019	Q28019 bos taurus
39	45	41.7	132	16 Q8TP91	Q8TP91 anabaena sp
40	45	41.7	210	16 Q9X9Z1	Q9X9Z1 streptomyce
41	45	41.7	233	4 Q9N0E2	Q9N0E2 homo sapien
42	45	41.7	337	10 Q8S0P8	Q8S0P8 oryza sativ
43	45	41.7	346	4 Q9NWY4	Q9NWY4 homo sapien
44	45	41.7	359	10 Q9AV86	Q9AV86 oryza sativ
45	45	41.7	397	16 Q9KY75	Q9KY75 streptomyce

ALIGNMENTS

RESULT 1
Q9YCW4 ID Q9YCW4 PRELIMINARY; PRT; 125 AA.
AC Q9YCW4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein APE1148.
GN APE1148.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcaeae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=9310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankal A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA80133.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 125 AA; 13011 MW; 5EA3CF9F929F1B21 CRC64;

Query Match 46.3%; Score 50; DB 17; Length 125;
Best Local Similarity 54.5%; Pred. No. 5.2;
Matches 12; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 PLPPEMSG--SLKRRPQGGP 20

Db 10 PLLPQDDGRPLLRRPQGGGP 31

[illegible]

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler K., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celisner S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003692; AAF54681.1; -
 DR EMBL; AY070944; AAL48566.1; -
 DR FlyBase; FBgn0037934; CG6830.
 DR InterPro; IPR004119; DUF227.
 DR Pfam; PF02958; DUF227.2.
 SQ SEQUENCE 891 AA; 101884 MW; B723623AB4561563 CRC64;

Query Match 45.4%; Score 49; DB 5; Length 891;
 Best Local Similarity 50.0%; Pred. No. 48;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LPPPMGSLKRRPQGE 18
 ||||| :| :|||
 Db 460 PLPELLAALKPQGPQE 477

RESULT 5
 ID Q8WQ61 PRELIMINARY; PRT; 1253 AA.
 AC Q8WQ61;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Eps-15 protein.
 GN Eps-15.
 OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Waikar Y., Robinson I.M., O'Kane C.J.;
 RT "Characterisation of Drosophila Eps15 gene."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ421624; CAD13484.1; -
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR000261; EPS15_repeat.
 DR Pfam; PF00036; efhand; 4.
 DR SMART; SM00054; EFh; 4.
 DR SMART; SM00027; EH; 3.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN.2.
 SQ SEQUENCE 1253 AA; 134440 MW; 54CC8C62BEF8FEBCB CRC64;

Query Match 45.4%; Score 49; DB 5; Length 1253;
 Best Local Similarity 44.4%; Pred. No. 67;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 LPPPMGSLKRRPQGE 19
 :|| :| :| :|||
 Db 228 MPPPPAGAMPAPSGEG 245

RESULT 6
 ID O95875 PRELIMINARY; PRT; 2157 AA.
 AC O95875;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE BAT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
 RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
 RA Lasky S., Hood L.;
 RT "Sequence of the human major histocompatibility complex class III
 region."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF129756; AAD18086.1; -
 SQ SEQUENCE 2157 AA; 228869 MW; 21B817F5B699B0DB CRC64;

Query Match 45.4%; Score 49; DB 4; Length 2157;
 Best Local Similarity 47.4%; Pred. No. 1.1e+02;
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPPPMGSLKRRPQGE 20
 ||| | | :|||
 Db 1576 LPPHSSGFLGSKPEGPGP 1594

RESULT 7
 ID Q96QC6 PRELIMINARY; PRT; 2157 AA.
 AC Q96QC6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE BAT2 protein.
 GN BAT2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

Wed Apr 16 07:59:30 2003

us-09-917-372-1_copy_216_235.rspt

```

RP SEQUENCE FROM N.A.
RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shina S., Tamiya G., Oka A., Inoko H.;
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP000505; BAB03391.1;
SQ SEQUENCE 2157 AA; 228858 MW; C500E03B15382C9 CRC64;

Query Match 45.4%; Score 49; DB 4; Length 2157;
Best Local Similarity 47.4%; Pred. No. 1.1e+02;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQGE 20
||| | | :|:| |
Db 1576 LPPEMSGSLKRRPQGE 1594

RESULT 8
Q925G9 PRELIMINARY; PRT; 930 AA.
AC Q925G9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE PDZ-RGS3 protein.
GN RGS3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-21197945; PubMed-11301003;
Lu Q., Sun E.E., Klein R.S., Flanagan J.G.;
RT "Ephrin-b reverse signaling is mediated by a novel PDZ-RGS protein and
selectively inhibits G protein-coupled chemoattraction.";
RL Cell 105:69-79(2001).
DR EMBL; AF350047; AAK38878.1;
DR MGD; MGI:1354734; Rgs3.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000342; RegI_Gprotein.
DR Pfam; PF00595; PDZ; 1.
DR Pfam; PF00615; RGS; 1.
DR ProDom; PD001580; Reg_of_prg; 1.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS0106; PDZ; 1.
DR PROSITE; PS0132; RGS; 1.
SQ SEQUENCE 930 AA; 102511 MW; FB0FCEA164D2F6A1 CRC64;

Query Match 44.4%; Score 48; DB 11; Length 930;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRPQGE 20
|:|:| | | | | | | |
Db 535 PVPESSASVOKRLPSQESP 554

RESULT 9
Q9ATZ5 PRELIMINARY; PRT; 1200 AA.
AC Q9ATZ5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Resistance-like protein P1-A.
GN P1-A.
OS Linum usitatissimum (Flax) (Linseed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC SEQUENCE FROM N.A.
RA MEDLINE-21096393; PubMed=11158537;
Dadds P.N., Lawrence G.J., Ellis J.G.;
RT "Six Amino Acid Changes Confined to the Leucine-Rich Repeat beta-
strand/beta-Turn Motif Determine the Difference between the P and P2
Rust Resistance Specificities in Flax.";
RL Plant Cell 13:163-178(2001).
DR EMBL; AF310958; AAK28803.1;
DR InterPro; IPR000767; Disease_resist.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR002182; NB-ARC.
DR InterPro; IPR000157; TIR_domain.
DR Pfam; PF00560; LRR; 8.
DR Pfam; PF00931; NB-ARC; 1.
DR Pfam; PF01582; TIR; 1.
DR PRINTS; PR00364; DISEASERSIST.
DR SMART; SM00255; TIR; 1.
SQ SEQUENCE 1200 AA; 136588 MW; 09739254B3117E05 CRC64;

Query Match 44.0%; Score 47.5; DB 10; Length 1200;
Best Local Similarity 55.0%; Pred. No. 1.1e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 PLPPEMSGSLKRRPQGE 19
||||: ||| | | | |
Db 741 PLPPKLDKSLKLRVMOGLG 760

RESULT 10
Q9MA81 PRELIMINARY; PRT; 618 AA.
AC Q9MA81;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE T12H1.35 protein (Fragment).
GN T12H1.35.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. COLUMBIA;
Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC 112H1 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SURCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL; AC009177; AAF27044.1;
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00249; myb_DNA_binding; 1.
DR SMART; SM00395; SANT; 1.
DR DNA-binding; Nuclear protein.
FT NON_TER 618
SQ SEQUENCE 618 AA; 68924 MW; 234EE798627BA164 CRC64;

Query Match 43.5%; Score 47; DB 10; Length 618;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQ 16
| | | | | | | | | |
Db 476 LQPTSGSLKRRPK 490

```

```

RESULT 11
Q8RXZ0
ID Q8RXZ0 PRELIMINARY; PRT; 820 AA.
AC Q8RXZ0
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 91.8 kDa protein (Fragment).
GN AT3G05380.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kaniya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Full Length cDNA Clones.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY080604; AAL86288.1; -.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 820 AA; 91757 MW; 98267CA802334008 CRC64;

Query Match 43.5%; Score 47; DB 10; Length 820;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LPPEMSGSLRRPQ 16
I | : ||||| :| :
Db 241 LQPQTSGSLRRKPK 255

RESULT 12
Q9MA51
ID Q9MA51 PRELIMINARY; PRT; 1055 AA.
AC Q9MA51;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE F22F7.18 protein.
GN F22F7.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F22F7 genomic sequence.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL; AC009606; AAF64544.1; -.
DR InterPro; IPR001005; MYB_DNA_Binding.
DR DR InterPro; IPR002999; Tudor.
DR Pfam; PF00249; myb_DNA-binding; 1.
DR SMART; SM00395; SANT; 1.
DR SMART; SM00333; TUDOR; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1055 AA; 118243 MW; F7B3A455BF5B1475 CRC64;

```

```

Query Match 43.5%; Score 47; DB 10; Length 1055;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LPPEMSGSLRRPQ 16
I | : ||||| :| :
Db 476 LQPQTSGSLRRKPK 490

RESULT 13
Q9VE07
ID Q9VE07 PRELIMINARY; PRT; 1542 AA.
AC Q9VE07;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CG6026 protein.
GN CG6026.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Balleg R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattie B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclev J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Weissman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003724; AAF55629.2; -.
DR FlyBase; FBgn0038676; CG6026.
DR InterPro; IPR002965; P-rich_extensn.
DR PRINTS; PR01217; PRICHTEXTENS.
SQ SEQUENCE 1542 AA; 168644 MW; 27C06E61A73FF016 CRC64;

Query Match 43.5%; Score 47; DB 5; Length 1542;
Best Local Similarity 52.9%; Pred. No. 1.6e+02;

```

Search completed: April 8, 2003, 10:51:39
Job time : 9.15513 secs

Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 4 PMSGSLKRRPQGE GP 20
| | | | | | | | | |
Db 1016 PYPDHMYPRPQAGP 1032

RESULT 14

Q95JD0 PRELIMINARY; PRT; 511 AA.
AC Q95JD0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Basic proline-rich protein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PAROTID GLAND;
RA Zhang Q., Szalay A.A., Kyeune-Nyombi E., Sands J.F., Oberg K.C.,
RA Tieche J.-M., Leonora J.;
RT "Cloning and expression of a novel proline-rich protein from porcine
parotid glands."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY035848; AAK61382.1;
DR PRINTS; PR01582; KV33CHANNEL.
SQ SEQUENCE 511 AA; 48483 MW; AB04597964C448D7 CRC64;

Query Match 43.1%; Score 46.5; DB 6; Length 511;
Best Local Similarity 47.8%; Pred. No. 68;
Matches 11; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 1 PLPPEMS---GSLKRRPQGE GP 20
| | | | | | | | | |
Db 46 PPPPEESQEGHQRPRPGDGP 68

RESULT 15

Q95JD1 PRELIMINARY; PRT; 566 AA.
AC Q95JD1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Basic proline-rich protein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PAROTID GLAND;
RA Zhang Q., Szalay A.A., Kyeune-Nyombi E., Sands J.F., Oberg K.C.,
RA Tieche J.-M., Leonora J.;
RT "Cloning and expression of a novel proline-rich protein from porcine
parotid glands."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY035847; AAK61381.1;
DR PRINTS; PR01582; KV33CHANNEL.
SQ SEQUENCE 566 AA; 53213 MW; E33B3B5E1BDEB81A CRC64;

Query Match 43.1%; Score 46.5; DB 6; Length 566;
Best Local Similarity 47.8%; Pred. No. 75;
Matches 11; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 1 PLPPEMS---GSLKRRPQGE GP 20
| | | | | | | | | |
Db 46 PPPPEESQEGHQRPRPGDGP 68

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM.protein - protein search, using sw model

Run on: April 8, 2003, 10:47:16 ; Search time 1.76611 seconds
(without alignments)
333.195 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPMSGSLKRRPQGECP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCRUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	42.6	197	US-08-505-606-1	Sequence 1, Appl
2	46	42.6	137	US-09-000-166-1	Sequence 1, Appl
3	46	42.6	1064	US-08-357-598-2	Sequence 2, Appl
4	46	42.6	1064	US-09-003-289-2	Sequence 2, Appl
5	46	42.6	1064	PCT-US95-16435-2	Sequence 2, Appl
6	46	42.6	1082	US-08-357-598-5	Sequence 5, Appl
7	46	42.6	1082	US-08-357-598-10	Sequence 10, Appl
8	46	42.6	1082	US-09-003-289-5	Sequence 5, Appl
9	46	42.6	1082	US-09-003-289-10	Sequence 10, Appl
10	46	42.6	1082	PCT-US95-16435-5	Sequence 5, Appl
11	46	42.6	1082	PCT-US95-16435-10	Sequence 10, Appl
12	46	42.6	1124	US-09-191-786-1	Sequence 1, Appl
13	46	42.6	1461	US-08-993-228-10	Sequence 10, Appl
14	45	41.7	415	US-09-006-353A-6	Sequence 6, Appl
15	45	41.7	415	US-09-573-986-6	Sequence 6, Appl
16	45	41.7	477	US-09-480-921B-10	Sequence 10, Appl
17	44	40.7	477	US-09-449-335-2	Sequence 2, Appl
18	44	40.7	477	US-09-449-335-6	Sequence 6, Appl
19	44	40.7	477	US-09-480-921B-29	Sequence 29, Appl
20	42	38.9	185	US-08-950-720A-13	Sequence 13, Appl
21	42	38.9	193	US-08-106-507-2	Sequence 2, Appl
22	42	38.9	302	US-08-893-853-3	Sequence 3, Appl
23	42	38.9	302	US-09-113-921-3	Sequence 3, Appl
24	42	38.9	894	US-08-117-362-4	Sequence 4, Appl
25	42	38.9	894	US-08-486-924-4	Sequence 4, Appl
26	42	38.9	1098	US-08-946-994-17	Sequence 17, Appl
27	42	38.9	1099	US-08-665-574C-16	Sequence 16, Appl

28	42	38.9	1100	1	US-08-357-598-11	Sequence 11, Appl
29	42	38.9	1100	2	US-09-003-289-11	Sequence 11, Appl
30	42	38.9	1100	5	PCT-US95-16435-11	Sequence 11, Appl
31	42	38.9	1299	5	PCT-US95-08354A-2	Sequence 2, Appl
32	42	38.9	1958	1	US-07-945-283-2	Sequence 2, Appl
33	41	38.0	563	4	US-09-422-936-79	Sequence 79, Appl
34	41	38.0	716	2	US-08-766-982-1	Sequence 1, Appl
35	41	38.0	716	4	US-09-296-219-1	Sequence 1, Appl
36	41	38.0	844	4	US-09-422-936-51	Sequence 51, Appl
37	41	38.0	886	4	US-09-422-936-77	Sequence 77, Appl
38	41	38.0	892	4	US-09-422-936-75	Sequence 75, Appl
39	41	38.0	899	4	US-09-422-936-71	Sequence 71, Appl
40	41	38.0	961	4	US-09-422-936-49	Sequence 49, Appl
41	40.5	37.5	1036	4	US-08-891-640-3	Sequence 3, Appl
42	40.5	37.5	1061	4	US-08-701-154A-5	Sequence 5, Appl
43	40.5	37.5	1336	2	US-08-231-193A-58	Sequence 58, Appl
44	40.5	37.5	1336	2	US-08-486-273A-58	Sequence 58, Appl
45	40.5	37.5	1336	3	US-08-940-086A-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-08-505-606-1
; Sequence 1, Application US/08505606
; Patent No. 5925351
; GENERAL INFORMATION:
; APPLICANT: BROWNING, Jeffrey L.
; APPLICANT: BENJAMIN, Christopher D.
; APPLICANT: HOCHMAN, Paula S.
; TITLE OF INVENTION: SOLUBLE LYMPHOTOXIN-BETA RECEPTORS AND
; TITLE OF INVENTION: ANTI-LYMPHOTOXIN RECEPTOR AND LIGAND ANTIBODIES AS
; TITLE OF INVENTION: THERAPEUTIC AGENTS FOR THE TREATMENT OF IMMUNOLOGICAL
; TITLE OF INVENTION: DISEASE
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr.
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,606
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378,968
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-505-606-1

Query Match 42.6%; Score 46; DB 2; Length 197;

```
Best Local Similarity 88.9%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLPPMSGSG 9
    |||||
Db 189 PLPPMSGSGT 197

RESULT 2
US-09-000-166-1
; Sequence 1, Application US/09000166A
; Patent No. 6403087
; GENERAL INFORMATION:
; APPLICANT: Browning, et al.
; TITLE OF INVENTION: Soluble Lymphotoxin-B Receptors and Anti-Lymphotoxin
; TITLE OF INVENTION: Receptor and Ligand Antibodies, as Therapeutic Agents
; TITLE OF INVENTION: for the Treatment of Immunological Disease.
; FILE REFERENCE: B191
; CURRENT APPLICATION NUMBER: US/09/000,166A
; CURRENT FILING DATE: 1998-06-08
; EARLIER APPLICATION NUMBER: PCT/US96/12010
; EARLIER FILING DATE: 1996-07-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-000-166-1

Query Match 42.6%; Score 46; DB 4; Length 197;
Best Local Similarity 88.9%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLPPMSGSG 9
    |||||
Db 189 PLPPMSGSGT 197

RESULT 3
US-08-357-598-2
; Sequence 2, Application US/08357598
; Patent No. 5705625
; GENERAL INFORMATION:
; APPLICANT: Civin, Curt I.
; APPLICANT: Small, Donald
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 15-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/033001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1064 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-357-598-2

Query Match 42.6%; Score 46; DB 2; Length 1064;
Best Local Similarity 47.1%; Pred. No. 71;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKRRPQG 17
    |||:|:|:|:|:|
Db 419 PLGPDYKGLIRRSPTG 435

RESULT 5
PCT-US95-16435-2
; Sequence 2, Application PC/TUS9516435
; GENERAL INFORMATION:
; APPLICANT: The Johns Hopkins University School of Medicine
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
```

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16435
FILING DATE: 15-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1064 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-16435-2

Query Match 42.6%; Score 46; DB 5; Length 1064;
Best Local Similarity 47.1%; Pred. No. 71;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKRRPQG 17
|||:|:|:|:|
Db 419 PLGPDYKGLIRSRPTG 435

RESULT 6
US-08-357-598-5
Sequence 5, Application US/08357598
Patent No. 5705625
GENERAL INFORMATION:
APPLICANT: Civin, Curt I.
APPLICANT: Small, Donald
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,598
FILING DATE: 15-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1082 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-357-598-5

Query Match 42.6%; Score 46; DB 1; Length 1082;
Best Local Similarity 47.1%; Pred. No. 72;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKRRPQG 17
|||:|:|:|:|
Db 419 PLGPDYKGLIRSRPTG 435

RESULT 7
US-08-357-598-10
Sequence 10, Application US/08357598
Patent No. 5705625
GENERAL INFORMATION:
APPLICANT: Civin, Curt I.
APPLICANT: Small, Donald
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,598
FILING DATE: 15-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1082 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-357-598-10

Query Match 42.6%; Score 46; DB 1; Length 1082;
Best Local Similarity 47.1%; Pred. No. 72;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKRRPQG 17
|||:|:~|:~|:~|
Db 419 PLGPDYKGLIRSRPTG 435

RESULT 8
US-09-003-289-5
Sequence 5, Application US/09003289
Patent No. 5916792
GENERAL INFORMATION:

APPLICANT: Civin, Curt I.
APPLICANT: Small, Donald
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/003,289
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/357,598
FILING DATE: 15-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033001
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1082 amino acids
TYPE: amino acid
STRANDEDNESS: No. 5916792 Relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-003-289-5
Query Match 42.6%; Score 46; DB 2; Length 1082;
Best Local Similarity 47.1%; Pred. No. 72;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 1 PLPPEMSGSLKRRPQG 17
Db 419 PLGPDYKGLIRSPGTG 435
RESULT 9
PCT-US95-16435-10
Sequence 10, Application US/09003289
Patent No. 5916792
GENERAL INFORMATION:
APPLICANT: Civin, Curt I.
APPLICANT: Small, Donald
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/003,289
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/357,598
FILING DATE: 15-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033001
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1082 amino acids
TYPE: amino acid
STRANDEDNESS: No. 5916792 Relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-003-289-10
Query Match 42.6%; Score 46; DB 2; Length 1082;
Best Local Similarity 47.1%; Pred. No. 72;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 1 PLPPEMSGSLKRRPQG 17
Db 419 PLGPDYKGLIRSPGTG 435
RESULT 10
PCT-US95-16435-5
Sequence 5, Application PC/TUS9516435
GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University School of Medicine
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16435
FILING DATE: 15-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/033W01
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1082 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-16435-5
Query Match 42.6%; Score 46; DB 5; Length 1082;
Best Local Similarity 47.1%; Pred. No. 72;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 1 PLPPEMSGSLKRRPQG 17
Db 419 PLGPDYKGLIRSPGTG 435

Db 419 PLGPDYKGLIRRSPTG 435

RESULT 11

PCT-US95-16435-10

Sequence 10, Application PC/TUS9516435

GENERAL INFORMATION:

APPLICANT: The Johns Hopkins University School of Medicine

TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/16435

FILING DATE: 15-DEC-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07265/033W01

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1082 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-16435-10

Query Match 42.6%; Score 46; DB 5; Length 1082;

Best Local Similarity 47.1%; Pred. No. 72;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKKRRPQG 17

Db 419 PLGPDYKGLIRRSPTG 435

RESULT 12

US-09-191-786-1

Sequence 1, Application US/09191786

Patent No. 6372898

GENERAL INFORMATION:

APPLICANT: Cacalano, Nicholas A.

APPLICANT: Johnston, James A.

TITLE OF INVENTION: Mammalian Protein Variants and Methods

TITLE OF INVENTION: of Use

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: DNAX Research Institute

STREET: 901 California Avenue

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/993,228

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.

REGISTRATION NUMBER: 41,323

TELEPHONE: (617) 498-8284

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1461 amino acids

QY 1 PLPPEMSGSLKKRRPQG 17

Db 420 PLGPDYKGLIRRSPTG 436

RESULT 13

US-08-993-228-10

Sequence 10, Application US/08993228

Patent No. 5976838

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

APPLICANT: McCoy, John M.

APPLICANT: Lavallie, Edward R.

APPLICANT: Racine, Lisa A.

APPLICANT: Merberg, David

APPLICANT: Treacy, Maurice

APPLICANT: Spaulding, Vikki

APPLICANT: Agostino, Michael J.

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

TITLE OF INVENTION: ENCODING THEM

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 CambridgePark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/993,228

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.

REGISTRATION NUMBER: 41,323

TELEPHONE: (617) 498-8284

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1461 amino acids

QY 1 PLPPEMSGSLKKRRPQG 17

Db 420 PLGPDYKGLIRRSPTG 436

Query Match 42.6%; Score 46; DB 4; Length 1124;

Best Local Similarity 47.1%; Pred. No. 75;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKKRRPQG 17

Db 420 PLGPDYKGLIRRSPTG 436

US-09-191-786-1

NAME: Ching, Edwin P.

REGISTRATION NUMBER: 34,090

REFERENCE/DOCKET NUMBER: DX0920

TELEPHONE: (650)852-9196

TELEFAX: (650)496-1200

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1124 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1062

US-09-191-786-1

; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-993-228-10

Query Match 42.6%; Score 46; DB 2; Length 1461;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 LPPEMSGSLKRRPQ 16
:||||| |
Db 272 VPDEMSADLEKRRPE 286

RESULT 14

US-09-006-353A-6
; Sequence 6, Application US/090006353A
; Patent No. 6261801

GENERAL INFORMATION:

; APPLICANT: WEI, YING-FEI
; APPLICANT: YU, GUO-LIANG
; APPLICANT: GENTZ, REINER
; APPLICANT: RUBEN, STEVEN
; TITLE OF INVENTION: TUMOR NECROSIS FACTOR RECEPTOR 5
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD

; COUNTRY: US

; ZIP: 20850

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,353A
; FILING DATE:

; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: BROOKES, ANDERS A
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF341
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 415 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-006-353A-6

Query Match 41.7%; Score 45; DB 4; Length 415;
Best Local Similarity 72.7%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 8 GSLLKRRPQGE 18
:||||| |
Db 255 GTLLKRRPQGE 265

RESULT 15

US-09-573-986-6

; Sequence 6, Application US/09573986

; Patent No. 6455040

GENERAL INFORMATION:

; APPLICANT: Wei, Ying-Fei

; APPLICANT: Ni, Jian
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488.1280004
; CURRENT APPLICATION NUMBER: US/09/573,986
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-573-986-6

Query Match 41.7%; Score 45; DB 4; Length 415;
Best Local Similarity 72.7%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 8 GSLLKRRPQGE 18
:||||| |
Db 255 GTLLKRRPQGE 265

Search completed: April 8, 2003, 10:53:21
Job time : 3.76611 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:51:46 ; Search time 2.14797 Seconds
(without alignments)
569.244 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPMSGSLKRRPQGECP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues
Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published_Applications_AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	108	100.0	399	10	US-09-907-372-1
2	80	74.1	257	10	US-09-948-018-19
3	80	74.1	435	10	US-09-907-372-19
4	80	74.1	435	10	US-09-768-779A-6
5	49	45.4	601	10	US-09-925-301-844
6	48	44.4	930	9	US-10-113-794A-1
7	46	42.6	197	9	US-10-003-211-1
8	46	42.6	348	9	US-09-789-054A-20
9	46	42.6	371	10	US-09-925-300-1681
10	46	42.6	1124	9	US-09-978-248-2
11	46	42.6	1124	10	US-09-771-161A-199
12	45	41.7	192	10	US-09-925-302-580
13	45	41.7	264	9	US-10-102-627-67
14	45	41.7	415	10	US-09-826-212-6
15	45	41.7	415	10	US-09-907-372-20
16	45	41.7	415	10	US-09-935-727-8
17	45	41.7	477	9	US-10-047-412A-10
18	45	41.7	543	10	US-09-927-738-2
19	44	40.7	235	9	US-10-050-704-102

20	44	40.7	244	9	US-10-050-704-209	Sequence 209, Appl
21	44	40.7	477	9	US-10-047-412A-29	Sequence 29, Appl
22	44	40.7	477	10	US-09-923-556-2	Sequence 2, Appl
23	44	40.7	477	10	US-09-923-556-6	Sequence 6, Appl
24	44	40.7	477	10	US-09-987-025-2	Sequence 2, Appl
25	44	40.7	726	10	US-09-770-689A-4	Sequence 4, Appl
26	44	40.7	803	10	US-09-770-689A-2	Sequence 2, Appl
27	44	40.7	881	10	US-09-816-860A-2	Sequence 2, Appl
28	42	38.9	172	9	US-09-738-626-4747	Sequence 4747, Ap
29	42	38.9	193	10	US-09-813-329-15	Sequence 15, Appl
30	42	38.9	255	9	US-09-738-626-3575	Sequence 3575, Ap
31	42	38.9	314	9	US-10-097-065-305	Sequence 305, App
32	42	38.9	418	10	US-09-771-161A-218	Sequence 218, App
33	42	38.9	428	9	US-10-125-635A-405	Sequence 405, App
34	42	38.9	570	10	US-09-877-633-13	Sequence 13, Appl
35	42	38.9	570	10	US-09-877-633-14	Sequence 14, Appl
36	42	38.9	1265	10	US-09-862-027-80	Sequence 80, Appl
37	41	38.0	53	10	US-09-864-761-39525	Sequence 39525, A
38	41	38.0	201	9	US-10-043-487-323	Sequence 323, App
39	41	38.0	345	9	US-09-738-626-3894	Sequence 3894, Ap
40	41	38.0	487	10	US-09-925-301-910	Sequence 910, App
41	41	38.0	527	9	US-10-022-832-6	Sequence 6, Appl
42	41	38.0	535	9	US-10-043-487-306	Sequence 306, App
43	41	38.0	793	10	US-09-875-724-4	Sequence 4, Appl
44	41	38.0	844	10	US-09-875-724-8	Sequence 8, Appl
45	41	38.0	960	10	US-09-826-508-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-907-372-1
; Sequence 1, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7497867CD1
US-09-907-372-1

Query Match 100.0%; Score 108; DB 10; Length 399;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 PLPPMSGSLKRRPQGECP 20
DB 216 PLPPMSGSLKRRPQGECP 235

RESULT 2
US-09-948-018-19
; Sequence 19, Application US/09948018
; Patent No. US20020150977A1
; GENERAL INFORMATION:
; APPLICANT: Theill et al
; TITLE OF INVENTION: TNF RECEPTOR-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCE: 01017737677
; CURRENT APPLICATION NUMBER: US/09/948,018
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,191

```
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-948-018-19

Query Match          74.1%; Score 80; DB 10; Length 257;
Best Local Similarity 35.7%; Pred. No. 0.00029;
Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 PLPPMS-----GSLKRRPQGECP 20
      |||||
Db 127 PLPPMSGTMLMLAVLLPLAFFLLLATVFCIWKSHPSLCRKLGSLKRRPQGECP 182

RESULT 3
US-09-907-372-19
; Sequence 19, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 19
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 g339762
US-09-907-372-19

Query Match          74.1%; Score 80; DB 10; Length 435;
Best Local Similarity 35.7%; Pred. No. 0.00052;
Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 PLPPMS-----GSLKRRPQGECP 20
      |||||
Db 216 PLPPMSGTMLMLAVLLPLAFFLLLATVFCIWKSHPSLCRKLGSLKRRPQGECP 271

RESULT 4
US-09-768-779A-6
; Sequence 6, Application US/09768779A
; Patent No. US20020127637A1
; GENERAL INFORMATION:
; APPLICANT: NI, JIAN
; MOORE, PAUL
; TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR
; RECEPTOR-LIKE PROTEIN 8
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/768,779A
```

```
; FILING DATE: 25-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/086,582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: KENLEY K. HOOVER
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PF368PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-768-779A-6

Query Match          74.1%; Score 80; DB 10; Length 435;
Best Local Similarity 35.7%; Pred. No. 0.00052;
Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 PLPPMS-----GSLKRRPQGECP 20
      |||||
Db 216 PLPPMSGTMLMLAVLLPLAFFLLLATVFCIWKSHPSLCRKLGSLKRRPQGECP 271

RESULT 5
US-09-925-301-844
; Sequence 844, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 844
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (64)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (103)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (106)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (152)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (358)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (383)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

US-09-925-301-844

Query Match 45.4%; Score 49; DB 10; Length 601;
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQGECP 20
||| | | :||| |
Db 194 LPPHSSGFLGSKPEGPG 212

RESULT 6

US-10-113-794A-1
; Sequence 1, Application US/10113794A
; Publication No. US20030022202A1
; GENERAL INFORMATION:
; APPLICANT: Flanagan et al.
; TITLE OF INVENTION: B EPHRIN REGULATION OF G-PROTEIN COUPLED
; TITLE OF INVENTION: CHEMOATTRACTION
; FILE REFERENCE: 2535/106
; CURRENT APPLICATION NUMBER: US/10/113,794A
; CURRENT FILING DATE: 2002-04-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-113-794A-1

Query Match 44.4%; Score 48; DB 9; Length 930;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 PLPEMSGSLKRRPQGECP 20
| : ||| : ||| | |
Db 535 VPPESSASVQKRLPSQESP 554

RESULT 7

US-10-003-211-1
; Sequence 1, Application US/10003211
; Publication No. US20020197254A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; TITLE OF INVENTION: Soluble Lymphotoxin Beta Receptor and
; TITLE OF INVENTION: Anti-Lymphotoxin Receptor and Ligand Antibodies as
; TITLE OF INVENTION: Therapeutic Agents for the Treatment of Immunological
; TITLE OF INVENTION: Diseases
; FILE REFERENCE: A013US
; CURRENT APPLICATION NUMBER: US/10/003,211
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: PCT/US97/19436
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/029,060
; PRIOR FILING DATE: 1996-10-25
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-003-211-1

Query Match 42.6%; Score 46; DB 9; Length 197;
Best Local Similarity 88.9%; Pred. No. 20;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLPEMSGS 9
| | | | | | | | | |
Db 189 PLPEMSGT 197

RESULT 8

US-09-789-054A-20
; Sequence 20, Application US/09789054A
; Publication No. US20020184659A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Zhu, Qun
; TITLE OF INVENTION: PLANT GENES ENCODING DRI AND DRAP1, A GLOBAL REPRESSOR COMPLEX
; TITLE OF INVENTION: TRANSCRIPTION
; FILE REFERENCE: BB107 US CIP
; CURRENT APPLICATION NUMBER: US/09/789,054A
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/485558
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/US98/16688
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/055,865
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-789-054A-20

Query Match 42.6%; Score 46; DB 9; Length 348;
Best Local Similarity 44.4%; Pred. No. 38;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 PPEMSGSLKRRPQGECP 20
|| | : |||| : |
Db 26 PPVPGGTRRRRPPQAQAP 43

RESULT 9

US-09-925-300-1681
; Sequence 1681, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1681
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1681

Query Match 42.6%; Score 46; DB 10; Length 371;
Best Local Similarity 52.6%; Pred. No. 41;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQGECP 20
||| ||| | : | : |
Db 271 LPPTSGSRTKQRLPGQKP 289

RESULT 10

US-09-978-248-2
; Sequence 2, Application US/09978248
; Publication No. US20020197690A1
; GENERAL INFORMATION:

; APPLICANT: Holland, Pamela M
; APPLICANT: Virca, Duke G
; APPLICANT: Bird, Timothy A
; APPLICANT: Garika, Kristen
; TITLE OF INVENTION: GID (GNK INTERACTING DECARBOXYLASE) AND METHODS OF USE
; FILE REFERENCE: 2499-1-001N
; CURRENT APPLICATION NUMBER: US/09/978,248
; CURRENT FILING DATE: 2001-10-16
; PRIOR FILING DATE: 2001-10-16
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 788
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-978-248-2

Query Match 42.6%; Score 46; DB 9; Length 788;
Best Local Similarity 52.6%; Pred. No. 96;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQGCP 20
Db 688 LPPTSGSRKQRLPGQRP 706

RESULT 11

US-09-771-161A-199
; Sequence 199, Application US/09771161A
; Patent No. US2002011081A1
; GENERAL INFORMATION:

; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR FILING DATE: 2000-11-28
; PRIOR FILING DATE: 2000-11-28
; PRIOR FILING DATE: 2000-06-15
; PRIOR FILING DATE: 2000-06-15
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 199
; LENGTH: 1124
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-199

Query Match 42.6%; Score 46; DB 10; Length 1124;
Best Local Similarity 47.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRPQG 17
Db 420 PLGPDYRGCLIRSPGT 436

RESULT 12

US-09-925-302-580
; Sequence 580, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P4104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR FILING DATE: 2000-03-08
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 580
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-580

Query Match 41.7%; Score 45; DB 10; Length 192;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 11; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

QY 2 LPPEMSGSLKRR---PQEGG 19
Db 162 LCPQVSGSIHKRKIHFFPQGWG 183

RESULT 13

US-10-102-627-67
; Sequence 67, Application US/10102627
; Publication No. US20030054377A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ24C1
; CURRENT APPLICATION NUMBER: US/10/102,627
; CURRENT FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 110
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-627-67

Query Match 41.7%; Score 45; DB 9; Length 264;
Best Local Similarity 53.3%; Pred. No. 39;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRP 15
Db 223 PLPQDQTESLLRRQP 237

RESULT 14

US-09-826-212-6
; Sequence 6, Application US/09826212
; Patent No. US20010021516A1
; GENERAL INFORMATION:
; APPLICANT: Wei, Ying-Fei
; APPLICANT: Gentz, Reiner
; APPLICANT: Ruben, Steven
; APPLICANT: Ni, Jian
; TITLE OF INVENTION: Tumor Necrosis Factor Receptor 5
; FILE REFERENCE: 1488.1280006
; CURRENT APPLICATION NUMBER: US/09/826,212
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-212-6

Query Match 41.7%; Score 45; DB 10; Length 415;
Best Local Similarity 72.7%; Pred. No. 65;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 8 GSLKRRPGE 18
|:||||| |:|
Db 255 GTLLKRHPEGE 265

RESULT 15
US-09-907-372-20
; Sequence 20, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 g600223
US-09-907-372-20

Query Match 41.7%; Score 45; DB 10; Length 415;
Best Local Similarity 72.7%; Pred. No. 65;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 GSLKRRPGE 18
|:||||| |:|
Db 255 GTLLKRHPEGE 265

Search completed: April 8, 2003, 11:03:51
Job time : 3.14797 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:05:15 ; Search time 344.98 seconds
(without alignments)
12938.300 Million cell updates/sec

Title: US-09-917-372-2
Perfect score: 1982
Sequence: 1 gccccgcgcagctcgtc.....ggtatggggaagagtttgg 1982

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1809.2	91.3	2136	24	Gene #2135 used to
2	1809.2	91.3	2136	24	Human benign prost
3	1809.2	91.3	2136	24	Breast cancer rela
4	1809.2	91.3	2136	24	Stomach cancer rel
5	1173.4	59.2	1594	24	Human ovarian anti
6	444.6	22.4	527	24	Human colon cancer
7	433.6	21.9	481	24	Human ovarian can
8	386.4	19.5	529	24	Human ovarian can
9	382	19.3	516	24	Human ovarian can

c	10	253.8	12.8	281	24	ABL82623
	11	251.2	12.7	289	24	ABL82688
	12	72	3.6	1557	15	AAQ45225
	13	72	3.6	1557	22	AAQ90446
	14	72	3.6	1640	12	AAQ10990
	15	72	3.6	1641	14	AAQ49931
	16	72	3.6	2393	12	AAQ10907
	17	72	3.6	2394	22	AAQ83951
	18	72	3.6	3683	24	ABK83997
	19	72	3.6	3683	24	ABK65877
	20	72	3.6	3683	24	ABK3465
	21	72	3.6	3683	24	ABK3466
	22	72	3.6	3683	24	ABK3467
	23	70.4	3.5	2224	16	AAQ89544
	24	69.2	3.5	1641	15	AAQ45224
	25	62.6	3.2	691	16	AAQ05443
	26	62.6	3.2	705	19	AAV41549
	27	62.6	3.2	705	19	AAV19802
	28	62.6	3.2	705	20	AAV81733
	29	62.6	3.2	705	21	AAZ45759
	30	62.6	3.2	705	21	AAZ45760
	31	62.6	3.2	705	21	AAZ45761
	32	62.6	3.2	705	21	AAZ45762
	33	62.4	3.1	927	19	AAV44852
	34	62.4	3.1	927	20	AAZ87265
	35	62.4	3.1	1049	21	AAA46931
	36	62.4	3.1	1049	21	AAZ49727
	37	62.4	3.1	1596	21	AAZ94196
	38	62.4	3.1	1704	21	AAA28135
	39	62.4	3.1	1724	18	AAZ51737
	40	62.4	3.1	1724	20	AAZ25512
	41	62.4	3.1	1724	24	ABK84410
	42	62.4	3.1	1815	22	AAZ03021
	43	62.4	3.1	1834	21	AAZ94198
	44	62.4	3.1	1929	21	AAZ94195
	45	62.4	3.1	2313	21	AAZ94197

ALIGNMENTS

RESULT 1	ABN95637
ID	ABN95637 standard; DNA; 2136 BP.
XX	ABN95637;
AC	ABN95637;
XX	
DT	13-AUG-2002 (first entry)
XX	
DE	Gene #2135 used to diagnose liver cancer.
XX	
KW	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW	metastatic liver tumour; cytostatic; expression profile; disease state;
KW	disease progression; drug toxicity; drug efficacy; drug metabolism.
XX	
OS	Homo sapiens.
XX	
PN	WO200229103-A2.
XX	
PD	11-APR-2002.
XX	
PF	02-OCT-2001; 2001WO-US30589.
XX	
PR	02-OCT-2000; 2000US-237054P.
XX	(GENE-) GENE LOGIC INC.
PA	
PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX	
DR	WPI; 2002-426119/45.
XX	
PT	Diagnosing and detecting the progression of liver cancer,
PT	hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a
 PS liver tissue sample
 XX Claim 1; SEQ ID NO 2135; 298pp; English.
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 91.3%; Score 1809.2; DB 24; Length 2136;
 Best Local Similarity 94.6%; Pred. No. 0;
 Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

QY 51 GCCCTGAGAGCGCGCTGCGCGCTCCCGGCTGGGTGCACATCGCCCTGAGTCCCG 110
 DB 1 GCCCTGAGAGCGCGCTGCGCGCTCCCGGCTGGGTGCACATCGCCCTGAGTCCCG 60
 QY 111 TCCAGGCTCTGGGCTGGGAGCGCGCGCGCTGCGCGCTGGGAGCGCTGCGGCTCTGCG 170
 DB 61 TCCAGGCTCTGGGCTGGGAGCGCGCGCGCTGCGCGCTGGGAGCGCTGCGGCTCTGCG 120
 QY 171 CTTCCTCCAGGCGCGCGCTGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGCTCTGCGCT 230
 DB 121 CTTCCTCCAGGCGCGCGCTGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGCTCTGCGCT 180
 QY 231 TGGGCGACCTCTGCGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGCTCTGCGGCTC 290
 DB 181 TGGGCGACCTCTGCGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGCTCTGCGGCTC 240
 QY 291 CTGGCAGCATCGAGCGCGCGCTGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGCTCTGCGG 350
 DB 241 CTGGCAGCATCGAGCGCGCGCTGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGCTCTGCGG 300
 QY 351 GACCAGAAAGGAATATATAGCGCGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGCTC 410
 DB 301 GACCAGAAAGGAATATATAGCGCGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGCTC 360
 QY 411 GGCACCTATGCTCAGCTAATGAGCGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGCTC 470
 DB 361 GGCACCTATGCTCAGCTAATGAGCGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGCTC 420
 QY 471 GAGAAATCTTACACAGAGCACTGAACTACTGACCATCTGCCAGCTGTGCGGCGCGCTGT 530
 DB 421 GAGAAATCTTACACAGAGCACTGAACTACTGACCATCTGCCAGCTGTGCGGCGCGCTGT 480
 QY 531 GACCCAGTGTAGGCGCTGAGGAGATGCGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGG 590
 DB 481 GACCCAGTGTAGGCGCTGAGGAGATGCGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGG 540
 QY 591 CGGTGCCAGCGGGAATGTTCTGTGCTGCTGGGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTA 650
 DB 541 CGGTGCCAGCGGGAATGTTCTGTGCTGCTGGGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTA 600
 QY 651 CTTCCTGACTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGAGGTTAAC 710
 DB 601 CTTCCTGACTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGAGGTTAAC 660
 QY 711 AACCACTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGAGGTTAAC 770
 DB 711 AACCACTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGAGGTTAAC 720

DB 661 AACCACTGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGAGGTTAAC 720
 QY 771 TGCAGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGAGGTTAAC 830
 DB 721 TGCAGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGAGGTTAAC 780
 QY 831 CAGTCCGACACACCTGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGAGGTTAAC 884
 DB 781 CAGTCCGACACACCTGCGCGCTGCGCGCTGCGCGCTGGGAGCGCTGCGGAGGTTAAC 840
 QY 885 ----- 884
 DB 841 ATGCTGATGCTGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 QY 885 -----GGATCGCTGCTCAAGAGG 902
 DB 901 TCCTGCTATCTGGAAGAGCGCGCGCTTCTCTGCGAGGAACTGGGATCGCTGCTCAAGAGG 960
 QY 903 CGTCCGAGGAGGAGCGCGCGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
 DB 961 CGTCCGAGGAGGAGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 QY 963 CCATACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
 DB 1021 CCATACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 QY 1023 TCCACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
 DB 1081 TCCACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 QY 1083 CTGGACCTGACAGGAGCGCGCTTGGAAACCGCGGAGCGAGCGCGCTGCGGCTGCGGCT 1142
 DB 1141 CTGGACCTGACAGGAGCGCGCTTGGAAACCGCGGAGCGAGCGCGCTGCGGCTGCGGCT 1200
 QY 1143 ACCAATGGCATTCATGTCACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
 DB 1201 ACCAATGGCATTCATGTCACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 QY 1203 AATGGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
 DB 1261 AATGGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
 QY 1263 CCATACCGCATTCGCGAGAGGCGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322
 DB 1321 CCATACCGCATTCGCGAGAGGCGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 QY 1323 GAAGATGGCAAGGCTTGGCACCTAGCGAGAGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382
 DB 1381 GAAGATGGCAAGGCTTGGCACCTAGCGAGAGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 QY 1383 AGGGGCGCAAGGCAACCAATTTATCACCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1442
 DB 1441 AGGGGCGCAAGGCAACCAATTTATCACCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 QY 1443 AGGGGCGCAAGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
 DB 1501 AGGGGCGCAAGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
 QY 1503 GCCTGAGTAGGCGCGCGGAGAGCGAGCGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCT 1562
 DB 1561 GCCTGAGTAGGCGCGCGGAGAGCGAGCGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCT 1620
 QY 1563 GCAGGTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1622
 DB 1621 GCAGGTGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
 QY 1623 ACCTGAGGCGCTGCGGAGAGCGCGCGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCT 1682
 DB 1681 ACCTGAGGCGCTGCGGAGAGCGCGCGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCT 1740
 QY 1683 AGGGCAGCATCATACCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1742
 DB 1741 AGGGCAGCATCATACCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800

QY 1743 GGAGCGCCACACGGTCACCTGCAAGGACGTACAGGCGCCCTCTAAAGGATTCGTGGTGC 1802
DB 1801 GGAGCGCCACACGGTCACCTGCAAGGACGTACAGGCGCCCTCTAAAGGATTCGTGGTGC 1860
QY 1803 TCATCCCCAAGCTTCAGAGACCCCTTTGGGGTTCCACACACTTCACGTGGACTGAGTAGACC 1862
DB 1861 TCATCCCCAAGCTTCAGAGACCCCTTTGGGGTTCCACACACTTCACGTGGACTGAGTAGACC 1920
QY 1863 CTGCATGAGATGAATATAGGAGGACGCTCTCTCCCTCCCTCTCTAGAGAGAGGAA 1922
DB 1921 CTGCATGAGATGAATATAGGAGGACGCTCTCTCCCTCCCTCTCTAGAGAGAGGAA 1980
QY 1923 AGGAGTTCATTAACTAGGAGGTTGGGTAGGATTCCTAGGTATGGGAAGAGTTTGG 1982
DB 1981 AGGAGTTCATTAACTAGGAGGTTGGGTAGGATTCCTAGGTATGGGAAGAGTTTGG 2040

RESULT 2
ABK64502
ID ABK64502 standard; DNA; 2136 BP.
AC ABK64502;
XX
XX 18-JUN-2002 (first entry)
XX Human benign prostatic hyperplasia gene #397.
XX Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX Homo sapiens.
XX WO200212440-A2.
XX 14-FEB-2002.
XX 07-AUG-2001; 2001WO-US24708.
XX 07-AUG-2000; 2000US-223323P.
XX 05-JUN-2001; 2001US-0873319.
XX (GENE-) GENE LOGIC INC.
XX (NISR) JAPAN TOBACCO INC.
XX Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
XX WPI; 2002-257476/30.
XX
XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by
XX detecting expression levels of one or more genes in prostate cells from
XX patient that are differentially regulated compared to normal prostate
XX cells -
XX
XX Disclosure; Page 229-230; 444pp; English.
XX
XX The invention relates to a method of diagnosing (I) the onset or
XX progression of benign prostatic hyperplasia (BPH), or screening (II) for
XX or identifying an agent that modulates the onset or progression of BPH.
XX The method is based on changes in gene expression in BPH tissue isolated
XX from patients exhibiting different clinical states of prostate
XX hyperplasia as compared to normal prostate tissue. (I) comprises
XX detecting the expression levels of one or more genes in prostate cells
XX from the subject that are differentially regulated compared to normal
XX prostate cells. (II) comprises preparing a first gene expression profile
XX of BPH cells or BPH-like cell population, exposing the cells to the
XX agent, preparing a second gene expression profile of the agent exposed
XX cells, and comparing the first and second gene expression profiles.
XX (I) is useful for diagnosing the onset or progression of BPH. (II) is
XX useful for identifying an agent that modulates the onset or progression
XX of BPH. The methods are useful to present information identifying
XX the expression level in a tissue or cells, by comparing the expression
XX level of genes given in the specification in the tissue or cells to the
XX level of expression of gene in the database, and displaying the

CC expression levels of at least one gene in the tissue or cell sample
CC compared to the expression level in BPH. Agents using (II) are useful for
CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
CC benign prostatic hyperplasia gene sequences of the invention.
XX
SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;
Query Match 91.3%; Score 1809.2; DB 24; Length 2136;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;
QY 51 GCCTTGAGGCGCGGCTGCGCGCTCCCGGCTTGGGGTGCACATCGGCCCTGAGTCCCG 110
DB 1 GCCTTGAGGCGCGGCTGCGCGCTCCCGGCTTGGGGTGCACATCGGCCCTGAGTCCCG 60
QY 111 TCCAGGCTCTGGGCTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTGC 170
DB 61 TCCAGGCTCTGGGCTGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTGC 120
QY 171 CTTCTCCCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 230
DB 121 CTTCTCCCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 180
QY 231 TGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 290
DB 181 TGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 240
QY 291 CTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 350
DB 241 CTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 300
QY 351 GACCAGGAAAGGAATACTATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 410
DB 301 GACCAGGAAAGGAATACTATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 411 GGCACCTATGCTCAGCTAAATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 470
DB 361 GGCACCTATGCTCAGCTAAATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 471 GAGAAATTCACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 530
DB 421 GAGAAATTCACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 531 GACCCAGTATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 590
DB 481 GACCCAGTATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 591 CGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTA 650
DB 541 CGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTA 600
QY 651 CTTTCTGACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTAAC 710
DB 601 CTTTCTGACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTAAC 660
QY 711 AACCACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 770
DB 661 AACCACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
QY 771 TGGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 830
DB 721 TGGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
QY 831 CAGTCCGACACACCTGCAAAAATCCATTAGAGCCACTGCCCGCCAGAGATGTCA----- 884
DB 781 CAGTCCGACACACCTGCAAAAATCCATTAGAGCCACTGCCCGCCAGAGATGTCAAGAAC 840
QY 885 ----- 884
DB 841 ATGCTGATGTCGCGGCTTCTGCTGCCACTGGCGCTCTTTCTGCTCTTCGACCGCTTC 900
QY 885 -----GGATCGCTGCTCAAGAGG 902
|||||

Db 901 TCCTGATCTGGAAGAGCCACCCCTTCTCTGCGAGAACTGGGATCGCTGCTCAAGAGG 960
QY 903 CGTCCGAGGAGAGAGGACCCCAATCCTGTAGCTGGAAGCTGGAGCCCTCCGGAAGCCCAT 962
Db 961 CGTCCGAGGAGAGAGGACCCCAATCCTGTAGCTGGAAGCTGGAGCCCTCCGGAAGCCCAT 1020
QY 963 CCATATCTCCCTGACTGTGTACAGCACTGCTACCACTTCTGGAGATGTTTCCCCAGTA 1022
Db 1021 CCATATCTCCCTGACTGTGTACAGCACTGCTACCACTTCTGGAGATGTTTCCCCAGTA 1080
QY 1023 TCCACTGGGCTCCCGGAGCCCACTTTTGGAGGAGGGGTGCGCAACAGCAGAGTCTCT 1082
Db 1081 TCCACTGGGCTCCCGGAGCCCACTTTTGGAGGAGGGGTGCGCAACAGCAGAGTCTCT 1140
QY 1083 CTGGACCTACACAGGAGCGGAGTTGGAACCCCGGGAGCAGAGCAGGTGCCACCGT 1142
Db 1141 CTGGACCTACACAGGAGCGGAGTTGGAACCCCGGGAGCAGAGCAGGTGCCACCGT 1200
QY 1143 ACCAATGCGATTCATGTCACCGGCGGGTCTATGACTATCATCTGCGCAACATCTACATCTAC 1202
Db 1201 ACCAATGCGATTCATGTCACCGGCGGGTCTATGACTATCATCTGCGCAACATCTACATCTAC 1260
QY 1203 AATGACCACTACTGGGGGACACCGGCTCTGTGAGACCTCCAGCTACCCCGGAACCT 1262
Db 1261 AATGACCACTACTGGGGGACACCGGCTCTGTGAGACCTCCAGCTACCCCGGAACCT 1320
QY 1263 CCATACCCCATTCGGAAGAGGGGACCTGCGCCCTCCGGGCTCTTACACCCCGACAG 1322
Db 1321 CCATACCCCATTCGGAAGAGGGGACCTGCGCCCTCCGGGCTCTTACACCCCGACAG 1380
QY 1323 GAAGATGCGAAGCTTGGCACTACCGAGACAGAGCACTGTGGTGCACACCCCTCTAAC 1382
Db 1381 GAAGATGCGAAGCTTGGCACTACCGAGACAGAGCACTGTGGTGCACACCCCTCTAAC 1440
QY 1383 AGGGGCCCAAGGAACCAATTTATCACCATGACTGACTGAGTCTCAGAAAGGCGAGAAGA 1442
Db 1441 AGGGGCCCAAGGAACCAATTTATCACCATGACTGACTGAGTCTCAGAAAGGCGAGAAGA 1500
QY 1443 AGGGGGGCACAGGACACTTCTCCCTTTAGGCTGCCCTGCCACGTGGGATTCACAGGG 1502
Db 1501 AGGGGGGCACAGGACACTTCTCCCTTTAGGCTGCCCTGCCACGTGGGATTCACAGGG 1560
QY 1503 GCCTGAGTAGGGCCCGGGAGCAGAGCCCTTAAGGGATTAAGGCTCAGACACCTCTGAGA 1562
Db 1561 GCCTGAGTAGGGCCCGGGAGCAGAGCCCTTAAGGGATTAAGGCTCAGACACCTCTGAGA 1620
QY 1563 GCAGTGGGCACTGGCTGGGTAGGTCGCTCCACAGGACTCTCCCTACTGCTGCTGAGCAA 1622
Db 1621 GCAGTGGGCACTGGCTGGGTAGGTCGCTCCACAGGACTCTCCCTACTGCTGCTGAGCAA 1680
QY 1623 ACCTGAGGCTCCCGGAGACCCACCCACCCCTGGGCTGCTCAGCCTCAGGCACGGAC 1682
Db 1681 ACCTGAGGCTCCCGGAGACCCACCCACCCCTGGGCTGCTCAGCCTCAGGCACGGAC 1740
QY 1683 AGGGACATGATACCACTGTGCCACTAGACAGCGCGGACCGGAGCAGCGCACCGAG 1742
Db 1741 AGGGACATGATACCACTGTGCCACTAGGACAGCGCGGACCGGAGCAGCGCACCGAG 1800
QY 1743 GGAGCGGCACACGGTCACTGCAAGGACGTCACGGGCGCCCTCTAAAGGATTCGTTGTC 1802
Db 1801 GGAGCGGCACACGGTCACTGCAAGGACGTCACGGGCGCCCTCTAAAGGATTCGTTGTC 1860
QY 1803 TCATCCCAAGCTTCAGAGACCCCTTTGGGGTTCCACACTTCACGTGGACTGAGGTAGACC 1862
Db 1861 TCATCCCAAGCTTCAGAGACCCCTTTGGGGTTCCACACTTCACGTGGACTGAGGTAGACC 1920
QY 1863 CTGCATGAAGTGAATTAATAGGAGGAGCGCTCTCCCTCCCTCCCTCTAGAGGAGAGAA 1922
Db 1921 CTGCATGAAGTGAATTAATAGGAGGAGCGCTCTCCCTCCCTCCCTCTAGAGGAGAGAA 1980
QY 1923 AGGGAGTCATTAAACACTAGGGGGTTGGGTAGGATTCCTAGGTATGGGAAGAGTTTGG 1982
Db 1981 AGGGAGTCATTAAACACTAGGGGGTTGGGTAGGATTCCTAGGTATGGGAAGAGTTTGG 2040

RESULT 3

ABL64078

ID ABL64078 standard; DNA; 2136 Bp.

XX ABL64078;

XX 15-MAY-2002 (first entry)

XX Breast cancer related gene sequence SEQ ID NO:2415.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
cystostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US10838.

XX 05-JUN-2000; 2000US-209473P.

XX 05-JUN-2000; 2000US-209531P.

XX 18-SEP-2000; 2000US-233133P.

XX 18-SEP-2000; 2000US-233617P.

XX 20-SEP-2000; 2000US-234009P.

XX 20-SEP-2000; 2000US-234034P.

XX 22-SEP-2000; 2000US-234052P.

XX 22-SEP-2000; 2000US-234509P.

XX 25-SEP-2000; 2000US-234567P.

XX 25-SEP-2000; 2000US-234923P.

XX 25-SEP-2000; 2000US-234924P.

XX 25-SEP-2000; 2000US-235077P.

XX 25-SEP-2000; 2000US-235082P.

XX 25-SEP-2000; 2000US-235134P.

XX 26-SEP-2000; 2000US-235280P.

XX 26-SEP-2000; 2000US-235637P.

XX 26-SEP-2000; 2000US-235638P.

XX 27-SEP-2000; 2000US-235711P.

XX 27-SEP-2000; 2000US-235720P.

XX 27-SEP-2000; 2000US-235840P.

XX 27-SEP-2000; 2000US-235863P.

XX 28-SEP-2000; 2000US-236028P.

XX 28-SEP-2000; 2000US-236032P.

XX 28-SEP-2000; 2000US-236033P.

XX 28-SEP-2000; 2000US-236034P.

XX 28-SEP-2000; 2000US-236109P.

XX 29-SEP-2000; 2000US-236111P.

XX 29-SEP-2000; 2000US-236842P.

XX 29-SEP-2000; 2000US-236891P.

XX 02-OCT-2000; 2000US-237172P.

XX 02-OCT-2000; 2000US-237173P.

XX 02-OCT-2000; 2000US-237278P.

XX 02-OCT-2000; 2000US-237294P.

XX 02-OCT-2000; 2000US-237295P.

XX 02-OCT-2000; 2000US-237316P.

XX 03-OCT-2000; 2000US-237425P.

XX 03-OCT-2000; 2000US-237598P.

XX 03-OCT-2000; 2000US-237604P.

XX 03-OCT-2000; 2000US-237606P.

XX 03-OCT-2000; 2000US-237608P.

XX 01-NOV-2000; 2000US-244867P.

XX 01-NOV-2000; 2000US-245084P.

XX (AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

XX Soppet DR, Weaver Z;

DR WPI: 2002-188264/24.
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX Claim 1; SEQ ID 2415; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 91.3%; Score 1809.2; DB 24; Length 2136;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;

QY 51 GCCTCGAGCCGGCTCGCCCTCCGGCCCTGGGTCACATCGGCCCTGAGTCCCG 110
DB 1 GCCTCGAGCCGGCTCGCCCTCCGGCCCTGGGTCACATCGGCCCTGAGTCCCG 60
QY 111 TCCAGGCTCTGGGCTCGGGCAGCCGCCACCGCTGCCAGAGCTGGGGCTCCTGC 170
DB 61 TCCAGGCTCTGGGCTCGGGCAGCCGCCACCGCTGCCAGAGCTGGGGCTCCTGC 120
QY 171 CTTCTCCAGGCCCCAGCTTCTGGCCCTGGCCCTGGCCAGTGGCCGCGATGCTTCGCT 230
DB 121 CTTCTCCAGGCCCCAGCTTCTGGCCCTGGCCCTGGCCAGTGGCCGCGATGCTTCGCT 180
QY 231 TGGCCACCTCTGCTCCCGCCCTGGCTGGGGCCCTGCTGGTGGCCCTTTCGGGCTC 290
DB 181 TGGCCACCTCTGCTCCCGCCCTGGCTGGGGCCCTGCTGGTGGCCCTTTCGGGCTC 240
QY 291 CTGGCAGATCGAGCCCGAGCGGTGCTTCATATGCTGCGGAGAACAGACTCGAG 350
DB 241 CTGGCAGATCGAGCCCGAGCGGTGCTTCATATGCTGCGGAGAACAGACTCGAG 300
QY 351 GACCCAGAAAGAAATATGAGCCCGAGCCAGCCGATCTGCTGCTCCGCTGCCGCCA 410
DB 301 GACCCAGAAAGAAATATGAGCCCGAGCCAGCCGATCTGCTGCTCCGCTGCCGCCA 360
QY 411 GGCACTATGCTCTAGCTAAATAGCCGCATCCGGGACACAGTTGTGTCACATGTGCC 470
DB 361 GGCACTATGCTCTAGCTAAATAGCCGCATCCGGGACACAGTTGTGTCACATGTGCC 420
QY 471 GAGAAATCTACAGCAGACTGGAATACCTGACCATCTGCCAGTGTGCCGCCCTGT 530
DB 421 GAGAAATCTACAGCAGACTGGAATACCTGACCATCTGCCAGTGTGCCGCCCTGT 480
QY 531 GACCCAGTATGGCCCTCGAGGAGATTGCCCTCTGCACAGCAACAGGAGCCAGTGC 590
DB 481 GACCCAGTATGGCCCTCGAGGAGATTGCCCTCTGCACAGCAACAGGAGCCAGTGC 540
QY 591 CGCTGCAGCCGGGAATGTCTGCTGCTGGCCCTCGAGTGTACACACTCGGAGCTA 650
DB 541 CGCTGCAGCCGGGAATGTCTGCTGCTGGCCCTCGAGTGTACACACTCGGAGCTA 600
QY 651 CTTTCTGACTGCCCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTAAC 710

DB 601 CTTTCTGACTGCCCGCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTAAC 660
QY 711 AACCACTCGCTCCCTCGAAGGAGGCACTTCCAGAAATACCTTCCCTCCCGGAGCCGCG 770
DB 661 AACCACTCGCTCCCTCGAAGGAGGCACTTCCAGAAATACCTTCCCTCCCGGAGCCGCG 720
QY 771 TGGCAGCCCCACACAGGCTGTGAGAACCAAGGCTTGGTGGAGGCACTCCAGGCACTGCC 830
DB 721 TGGCAGCCCCACACAGGCTGTGAGAACCAAGGCTTGGTGGAGGCACTCCAGGCACTGCC 780
QY 831 CAGTCCGACACAACTTGCAGAAATCCATTAGAGCCACTGCCGCCAGAGATGTCA----- 884
DB 781 CAGTCCGACACAACTTGCAGAAATCCATTAGAGCCACTGCCGCCAGAGATGTCAAGAAC 840
QY 885 ----- 884
DB 841 ATGCTGATGCTGCCCGCTTCTGCTGCCACTGGCCCTTCTTCTGCTCCTTGGCACCGTCTTC 900
QY 885 -----GGATCGCTGCTCAAGAGG 902
DB 901 TCTGCTGATGGAAGAGCCACCTTCTCTGTCAGGAAACTGGGATCGCTGCTCAAGAGG 960
QY 903 CGTCCCGAGGAGAGGAGCCCAATCCTGTAGCTGGAAGCTGGGAGCTCCGAAAGGCCAT 962
DB 961 CGTCCCGAGGAGAGGAGCCCAATCCTGTAGCTGGAAGCTGGGAGCTCCGAAAGGCCAT 1020
QY 963 CCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
DB 1021 CCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1023 TCCACATGGGCTCCCGCAGCCCGCTTGGAGGAGGCTGCCGAAACAGAGAGTCT 1082
DB 1081 TCCACATGGGCTCCCGCAGCCCGCTTGGAGGAGGCTGCCGAAACAGAGAGTCT 1140
QY 1083 CTGGACCTGACACAGGAGCGCGAGTTGGAAACCCGGGAGCAGAGGCTGGCCCGCAGGT 1142
DB 1141 CTGGACCTGACACAGGAGCGCGAGTTGGAAACCCGGGAGCAGAGGCTGGCCCGCAGGT 1200
QY 1143 ACCAATGGCATTCATGCTACCCGGGGTCTATGACTATCACTGGCAACATCTACATCTAC 1202
DB 1201 ACCAATGGCATTCATGCTACCCGGGGTCTATGACTATCACTGGCAACATCTACATCTAC 1260
QY 1203 AATGGACCACTGCTGGGGGACCCAGCGGTCTGAGAGCTCCAGCTACCCCGGAGACT 1262
DB 1261 AATGGACCACTGCTGGGGGACCCAGCGGTCTGAGAGCTCCAGCTACCCCGGAGACT 1320
QY 1263 CCATACCCCATTCGCCAGAGGAGGCGCTTGGCCCTCCCGGGCTCTCTACACCCCGCAG 1322
DB 1321 CCATACCCCATTCGCCAGAGGAGGCGCTTGGCCCTCCCGGGCTCTCTACACCCCGCAG 1380
QY 1323 GAAGATGGCAAGGCTTGGCACTAGCGGAGACAGAGCTGCTGGTGGCCACACCTCTAAC 1382
DB 1381 GAAGATGGCAAGGCTTGGCACTAGCGGAGACAGAGCTGCTGGTGGCCACACCTCTAAC 1440
QY 1383 AGGGGCCCAAGAACCAATTTATCAACCATGCTGCTGAGTCTGAGAAAGGCAAGA 1442
DB 1441 AGGGGCCCAAGAACCAATTTATCAACCATGCTGCTGAGTCTGAGAAAGGCAAGA 1500
QY 1443 AGGGGGCACAAGGAGGAGCTTCTCCCTTGGAGCTGCCCTGCCAGCTGGGATTCACAGGG 1502
DB 1501 AGGGGGCACAAGGAGGAGCTTCTCCCTTGGAGCTGCCCTGCCAGCTGGGATTCACAGGG 1560
QY 1503 GCCTGAGTAGGGCCCGGGAGAGAGCCCTAAGGATTAAGGCTCAGACACCTCTGAGA 1562
DB 1561 GCCTGAGTAGGGCCCGGGAGAGAGCCCTAAGGATTAAGGCTCAGACACCTCTGAGA 1620
QY 1563 GCAGGTGGGCACTGGGTGAGTGGCTTCCCTCCACAGGACTCTCCCTACTGCTGAGCAA 1622
DB 1621 GCAGGTGGGCACTGGGTGAGTGGCTTCCCTCCACAGGACTCTCCCTACTGCTGAGCAA 1680
QY 1623 ACCTGAGGCTCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1682

```

Db 1681 ACCTGAGGCTCCGGCAGACCCACCCACCCCTTGGGCTGCTCAGCCTCAGGCACCGAC 1740
QY 1683 AGGCACATGATACCACTGCTGCCACTACAGCAGCGCCGACCGGAGCAGCGACCGAG 1742
Db 1741 AGGCACATGATACCACTGCTGCCACTACGCGACCGCGGACCGGAGCAGCGACCGAG 1800
QY 1743 GGAGCGCCACACAGGTCACTGCAAGGACGTCACGGGCCCCCTCTAAAGGATTCGTGGTC 1802
Db 1801 GGAGCGCCACACAGGTCACTGCAAGGACGTCACGGGCCCCCTCTAAAGGATTCGTGGTC 1860
QY 1803 TCATCCCCAAGCTTCAGAGACCCCTTTGGGGTTCCACACTTCACGTGCACTAGTAGACC 1862
Db 1861 TCATCCCCAAGCTTCAGAGACCCCTTTGGGGTTCCACACTTCACGTGCACTAGTAGACC 1920
QY 1863 CTGCATGAGATGAATATATAGGAGGACGCTCCTCCCTCCCTCTAGAGAGAGAA 1922
Db 1921 CTGCATGAGATGAATATATAGGAGGACGCTCCTCCCTCCCTCTAGAGAGAGAA 1980
QY 1923 AGGAGTCATTAAACACTAGGGGTTGGGTAGGATTCCTAGGTATGGGAAGATTTTGG 1982
Db 1981 AGGAGTCATTAAACACTAGGGGTTGGGTAGGATTCCTAGGTATGGGAAGATTTTGG 2040

RESULT 4
ABL64384
ID ABL64384 standard; DNA; 2136 BP.
XX AC ABL64384;
XX DT 15-MAY-2002 (first entry)
XX DE Stomach cancer related gene sequence SEQ ID NO:2721.
XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
XX gene; ds.
XX OS Homo sapiens.
XX PN WO200194629-A2.
XX PD 13-DEC-2001.
XX PF 30-MAY-2001; 2001WO-US10838..
XX PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.

```

```

PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24..
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX Claim 1; SEQ ID 2721; 44pp; English.
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;
SQ
Query Match 91.3%; Score 1809.2; DB 24; Length 2136;
Best Local Similarity 94.6%; Pred. No. 0;
Matches 1929; Conservative 0; Mismatches 3; Indels 108; Gaps 1;
QY 51 GCCCTGGAGGCCGCGCTGGCGGCTCCCGGCCCTGGGGTGACATCGGCCCTGAGTCCCG 110
Db 1 GCCCTGGAGGCCGCGCTGGCGGCTCCCGGCCCTGGGGTGACATCGGCCCTGAGTCCCG 60
QY 111 TCCAGGCTCTGGGGTCTGGGGAGCGCGGCCACCGCTGCCAGGACGTGCGGCTCTGCG 170
Db 61 TCCAGGCTCTGGGGTCTGGGGAGCGCGGCCACCGCTGCCAGGACGTGCGGCTCTGCG 120
QY 171 CTTCCTCCAGGCCCGCCAGCTTGTGGCGGCTGGCCGAGTGGCGCCATGCTCTGCTGCT 230
Db 121 CTTCCTCCAGGCCCGCCAGCTTGTGGCGGCTGGCCGAGTGGCGCCATGCTCTGCTGCT 180
QY 231 TGGGCCACCTCTGCGCCCGCGGCTGGCCCTGGGGGCTCTGTGCTGCGGCTCTTGGGGTC 290
Db 181 TGGGCCACCTCTGCGCCCGCGGCTGGCCCTGGGGGCTCTGTGCTGCGGCTCTTGGGGTC 240
QY 291 CTGGCAGCATCGACGCCCGGCGGTGCTCCATATGCTCGGAGAACACACCTGCGAG 350
Db 241 CTGGCAGCATCGACGCCCGGCGGTGCTCCATATGCTCGGAGAACACACCTGCGAG 300

```

QY 351 GACCAGAAAGGAATACTATGAGCCCGACACCGCATCTGCTCCCGCTGCCGCGCA 410
Db 301 GACCAGAAAGGAATACTATGAGCCCGACACCGCATCTGCTCCCGCTGCCGCGCA 360
QY 411 GGCACCTATGCTCAGCTAAATGTACCGGATCCCGGACACAGTTTGTGCGCACATGTGC 470
Db 361 GGCACCTATGCTCAGCTAAATGTACCGGATCCCGGACACAGTTTGTGCGCACATGTGC 420
QY 471 GAGAATTCCTACAAAGGACCTGGAATACCTGACCATCTGCCAGCTGTGCCGCCCTGT 530
Db 421 GAGAATTCCTACAAAGGACCTGGAATACCTGACCATCTGCCAGCTGTGCCGCCCTGT 480
QY 531 GACCCAGTATGAGGCTCGAGGAGATTGCCCTCGACAAAGCAACCGAAGACCCAGTGC 590
Db 481 GACCCAGTATGAGGCTCGAGGAGATTGCCCTCGACAAAGCAACCGAAGACCCAGTGC 540
QY 591 CGTGCCAGCCGGGAATGTTCTGTGCTGCTGGGCCCTCGAGTGTACACATCGCGAGCTA 650
Db 541 CGTGCCAGCCGGGAATGTTCTGTGCTGCTGGGCCCTCGAGTGTACACATCGCGAGCTA 600
QY 651 CTTTCTGACTGCCCGCTGAGCGAGCTCAAGAGATGAAGTTGGGAAGGTTAAC 710
Db 601 CTTTCTGACTGCCCGCTGAGCGAGCTCAAGAGATGAAGTTGGGAAGGTTAAC 660
QY 711 AACCACTGCGTCCCTCGAAGGCGAGGCACTTCCAGAATACCTCTCCCGCAGCGCCGC 770
Db 661 AACCACTGCGTCCCTCGAAGGCGAGGCACTTCCAGAATACCTCTCCCGCAGCGCCGC 720
QY 771 TGCCAGCCCGACACAGGTGTGAGAACCAAGGTCTGGTGAGGAGCGCTCCAGCACATGCC 830
Db 721 TGCCAGCCCGACACAGGTGTGAGAACCAAGGTCTGGTGAGGAGCGCTCCAGCACATGCC 780
QY 831 CAGTCCGACACAACTGCAAAAATCCATTAGAGCCACTGCCCGCCAGAGATGTCA 884
Db 781 CAGTCCGACACAACTGCAAAAATCCATTAGAGCCACTGCCCGCCAGAGATGTCA 840
QY 885 ----- 884
Db 841 ATGCTGATGTGCGCGTTCTGCTGCCACTGGCGCTTCTTTCTGCTCTTGCCACCGTCTTC 900
QY 885 -----CGATCGCTGCTCAAGAG 902
Db 901 TCTGTGATGTGGAAGAGCCACCCTTCTCTGTGAGGAAATCGGGATCGCTCAAGAGG 960
QY 903 CGTCCCGAGGAGAGGCCAATCTGTAGCTGGAAGCTGGGAGCTCCGAAAGGCCAT 962
Db 961 CGTCCCGAGGAGAGGCCAATCTGTAGCTGGAAGCTGGGAGCTCCGAAAGGCCAT 1020
QY 963 CCATACTTCCCTGACTTGGGTACAGCCACTGCTACCCATTCTGGAGATGTTCCCGCAGTA 1022
Db 1021 CCATACTTCCCTGACTTGGGTACAGCCACTGCTACCCATTCTGGAGATGTTCCCGCAGTA 1080
QY 1023 TCACATGGGCTCCCGAGCCCGCAGTTTGGAGGAGGGGTGCCGCAACAGCAGTCTCT 1082
Db 1081 TCACATGGGCTCCCGAGCCCGCAGTTTGGAGGAGGGGTGCCGCAACAGCAGTCTCT 1140
QY 1083 CTGCACTGTACAGGAGCGCAGTTGGAAACCGGAGCAGAGCCAGGTGGCCCGCAGGT 1142
Db 1141 CTGCACTGTACAGGAGCGCAGTTGGAAACCGGAGCAGAGCCAGGTGGCCCGCAGGT 1200
QY 1143 ACCAATGGCATTCATGCTACCGCGGGTCTATGACTATCACTGGCAACATCACTAC 1202
Db 1201 ACCAATGGCATTCATGCTACCGCGGGTCTATGACTATCACTGGCAACATCACTAC 1260
QY 1203 AATGGACAGTACTGGGGGACACCGGGTCTGGAGACCTCCAGCTACCCCGCAACCT 1262
Db 1261 AATGGACAGTACTGGGGGACACCGGGTCTGGAGACCTCCAGCTACCCCGCAACCT 1320
QY 1263 CCATACCCCATTCGGAAGAGGGGACCTTGCGCCCTCCCGGCTCTCTACACCCCGCAG 1322
Db 1321 CCATACCCCATTCGGAAGAGGGGACCTTGCGCCCTCCCGGCTCTCTACACCCCGCAG 1380
QY 1323 GAAGATGGCAAGGCTTGGCACCTTAGCGGAGACAGACACTGTGTGTGCCACACCTCTAAC 1382

Db 1381 GAAGATGGCAAGGCTTGGCACCTTAGCGGAGACAGACACTGTGTGTGCCACACCTCTAAC 1440
QY 1383 AGGGGCCCCAAGGAACCAATTTATCACCCTACTGACTGACTGAGTCTGAGAAAAGCAGA 1442
Db 1441 AGGGGCCCCAAGGAACCAATTTATCACCCTACTGACTGAGGAGTCTGAGAAAAGCAGA 1500
QY 1443 AGGGGGGCAACAAGGCGACCTTCTCCCTTGAGGCTGCCCTGCCACCTGGGATTACAGGG 1502
Db 1501 AGGGGGGCAACAAGGCGACCTTCTCCCTTGAGGCTGCCCTGCCACCTGGGATTACAGGG 1560
QY 1503 GCCTGAGTAGGGCCCCGGGGAAGCAGAGCCCTTAAGGATTAAAGCTCAGACACCTCTGAGA 1562
Db 1561 GCCTGAGTAGGGCCCCGGGGAAGCAGAGCCCTTAAGGATTAAAGCTCAGACACCTCTGAGA 1620
QY 1563 GCAGGTGGGCACTGGCTGGGTACGGTGGCTCCACAGGACTCTCCCTACTGCTGAGCAA 1622
Db 1621 GCAGGTGGGCACTGGCTGGGTACGGTGGCTCCACAGGACTCTCCCTACTGCTGAGCAA 1680
QY 1623 ACCTGAGGCTCCCGGAGACCCACACCCCTGGGGCTGCTCAGCCTCAGGCACGGAC 1682
Db 1681 ACCTGAGGCTCCCGGAGACCCACACCCCTGGGGCTGCTCAGCCTCAGGCACGGAC 1740
QY 1683 AGGCGACATGATACCAACTGCTGCCACCTACAGCAGCCGCCACCGGAGCAGCAGCGAG 1742
Db 1741 AGGCGACATGATACCAACTGCTGCCACCTACAGCAGCCGCCACCGGAGCAGCAGCGAG 1800
QY 1743 GGAGCCGCCACACGGTCACCTGCAAGGACGTCACGGGCCCTCTAAGAGATTGCTGGTGC 1802
Db 1801 GGAGCCGCCACACGGTCACCTGCAAGGACGTCACGGGCCCTCTAAGAGATTGCTGGTGC 1860
QY 1803 TCATCCCCAAGCTTCAGAGACCTTTGGGTTCCACACTTCACACTGAGCTGAGGTAGACC 1862
Db 1861 TCATCCCCAAGCTTCAGAGACCTTTGGGTTCCACACTTCACACTGAGGTAGACC 1920
QY 1863 CTGCATGAAGATGAATATTATAGGAGAGCCTCTTCCCTCCCTCTCTAGAGGAGAGAA 1922
Db 1921 CTGCATGAAGATGAATATTATAGGAGAGCCTCTTCCCTCCCTCTCTAGAGGAGAGAA 1980
QY 1923 AGGAGTCATTACAACACTAGGGGGTTGGGTAGGATTCCTAGGTATGGGAGAGATTGG 1982
Db 1981 AGGAGTCATTACAACACTAGGGGGTTGGGTAGGATTCCTAGGTATGGGAGAGATTGG 2040

RESULT 5
ABQ55003
ID ABQ55003 standard; cDNA; 1594 BP.
XX
XX ABQ55003;
XX AC
XX DT
XX 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HSABJ44 cDNA, SEQ ID NO:883.
XX
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; chromosome 12p13;
KW gene; ss.
OS Homo sapiens.
XX
XX WO200200677-A1.
PN
XX
XX 03-JAN-2002.
PD
XX
XX 07-JUN-2001; 2001WO-US18569.
PF

XX 07-JUN-2000; 2000US-209467P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Birse CE, Rosen CA;
 XX WPI; 2002-147878/19.
 DR P-PSDB; ABP41926.
 XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 XX useful in the prevention, treatment and diagnosis of cancer (e.g.
 XX ovarian cancer), immune disorders, cardiovascular disorders and
 XX neurological diseases.
 XX Claim 1; SEQ ID No 883; 2922pp; English.
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 1594 BP; 391 A; 488 C; 434 G; 280 T; 1 other;
 SQ
 Query Match 59.2%; Score 1173.4; DB 24; Length 1594;
 Best Local Similarity 92.2%; Pred. NO. 4.7e-261;
 Matches 1292; Conservative 0; Mismatches 2; Indels 108; Gaps 1;
 QY 689 AGATGAAGTTGGAGGGTAACAACCACTGGCTCCCTGCAAGGAGGCGCACTTCCAGAA 748
 DB |||||||
 QY 91 AGATGAAGTTGGAGGGTAACAACCACTGGCTCCCTGCAAGGAGGNCATCTCCAGAA 150
 DB |||||||
 QY 749 TACCTCCTCCCGAGCGCCGCTGCGAGCCACACCACTAGGTGTGAGAACCAAGTCTGTGT 808
 DB |||||||
 DB 151 TACCTCCTCCCGAGCGCCGCTGCGAGCCACACCACTAGGTGTGAGAACCAAGTCTGTGT 210
 QY 809 GGAGGAGCTCCAGGCACTGCCAGTCCGACACAAACCTGCAAAAATCCATTAGAGCCACT 868
 DB |||||||
 DB 211 GGAGGAGCTCCAGGCACTGCCAGTCCGACACAAACCTGCAAAAATCCATTAGAGCCACT 270
 QY 869 GCGCCACAGATGTCA----- 884
 DB |||||||
 DB 271 GCGCCACAGATGTTCAGGAACCAATGCTGATGCTGCGCGTTCGTCGCACTGGCCTTCTT 330
 QY 885 ----- 884
 DB 331 TCTGCTCCTTGCACCGCTTCTCTGCTGATCTGGAGAGCCACCTTCTCTCTGAGGNA 390

QY 885 ----GGATCGCTGCTCAAGAGCGTCCGAGGAGGAGGAGCCCAATCTCTAGCTGGAAG 940
 DB |||||||
 DB 391 ACTGGGAGCTGCTCAAGAGCGTCCGAGGAGGAGGAGCCCAATCTCTAGCTGGAAG 450
 QY 941 CTGGGAGCTCCGAAAGGCCCATCATACTTCCCTGACTTGGTACAGCACTGCTACCCAT 1000
 DB |||||||
 DB 451 CTGGGAGCTCCGAAAGGCCCATCATACTTCCCTGACTTGGTACAGCACTGCTACCCAT 510
 QY 1001 TCTTGAGATGTTTCCCACTATCCACTGGGTCCCGGAGGCCAGTTTGGAGGAGG 1060
 DB |||||||
 DB 511 TCTTGAGATGTTTCCCACTATCCACTGGGTCCCGGAGGCCAGTTTGGAGGAGG 570
 QY 1061 GGTGCCCAACAGCAGAGTCTCTGGACCTGACAGGAGGCCGAGTTGGAACCCCGGGA 1120
 DB |||||||
 DB 571 GGTGCCCAACAGCAGAGTCTCTGGACCTGACAGGAGGCCGAGTTGGAACCCCGGGA 630
 QY 1121 GCAGAGCCAGGTGGCCCACTGACCAATGGCATTTGTCACCGCGGGTCTATGACTAT 1180
 DB |||||||
 DB 631 GCAGAGCCAGGTGGCCCACTGACCAATGGCATTTGTCACCGCGGGTCTATGACTAT 690
 QY 1181 CACTGGCAACATCTACATCTACATGACCACTGCTGGGGGACCCCGGTCTCTGAGA 1240
 DB |||||||
 DB 691 CACTGGCAACATCTACATCTACATGACCACTGCTGGGGGACCCCGGTCTCTGAGA 750
 QY 1241 CTTCCAGCTACCCCGAACCTCCATACCCCATTTCCGAAAGAGGGGACCTGGCCCTCC 1300
 DB |||||||
 DB 751 CTTCCAGCTACCCCGAACCTCCATACCCCATTTCCGAAAGAGGGGACCTGGCCCTCC 810
 QY 1301 CGGCTCTTACACCCACAGGAGATGCAAGGCTTGCACCTAGCGGAGCAGAGCA 1360
 DB |||||||
 DB 811 CGGCTCTTACACCCACAGGAGATGCAAGGCTTGCACCTAGCGGAGCAGAGCA 870
 QY 1361 CTGTGTGTCACACCTCTTAACAGGGGCCCAAGGAACCAATTTATCACCCTGACT 1420
 DB |||||||
 DB 871 CTGTGTGTCACACCTCTTAACAGGGGCCCAAGGAACCAATTTATCACCCTGACT 930
 QY 1421 GAGTCTGAGAAAGGCAAGAGAGGGGGGCACAAAGGACCTTCTCCTTGAAGCTGCC 1480
 DB |||||||
 DB 931 GAGTCTGAGAAAGGCAAGAGAGGGGGGCACAAAGGACCTTCTCCTTGAAGCTGCC 990
 QY 1481 TGCCCACTGGGATTCACAGGGGCTGAGTAGGGCCCGGGGAAGCAGAGCCCTAAGGAT 1540
 DB |||||||
 DB 991 TGCCCACTGGGATTCACAGGGGCTGAGTAGGGCCCGGGGAAGCAGAGCCCTAAGGAT 1050
 QY 1541 TAAGGCTCAGACACTCTGAGAGAGGTGGGCACTGGCTGGGTACGGTCCCTCCACAG 1600
 DB |||||||
 DB 1051 TAAGGCTCAGACACTCTGAGAGAGGTGGGCACTGGCTGGGTACGGTCCCTCCACAG 1110
 QY 1601 ACTCTCCCTACTGCTGAGCAAACTGAGGCTCCCGGAGAGCCACCCACCCCTGGG 1660
 DB |||||||
 DB 1111 ACTCTCCCTACTGCTGAGCAAACTGAGGCTCCCGGAGAGCCACCCACCCCTGGG 1170
 QY 1661 CTGCTCAGCCCTCAGGCAAGGACATGATACCAACTGCTGCCACTACAGCAGC 1720
 DB |||||||
 DB 1171 CTGCTCAGCCCTCAGGCAAGGACATGATACCAACTGCTGCCACTACGCGAGC 1230
 QY 1721 CGCACCGGAGCAGGCAAGGAGCGCCACAGGTCTACCTGCAAGGAGCTCAGGGC 1780
 DB |||||||
 DB 1231 CGCACCGGAGCAGGCAAGGAGCGCCACAGGTCTACCTGCAAGGAGCTCAGGGC 1290
 QY 1781 CCTCTAAAGATTGCTGCTCTATCCCAAGCTTCAGAGACCTTTGGGGTTCCACAC 1840
 DB |||||||
 DB 1291 CCTCTAAAGATTGCTGCTCTATCCCAAGCTTCAGAGACCTTTGGGGTTCCACAC 1350
 QY 1841 TTCAGTGGATGAGTACCCCTGCATGAAGATGAATATAGGAGGAGGCTCTTCC 1900
 DB |||||||
 DB 1351 TTCAGTGGATGAGTACCCCTGCATGAAGATGAATATAGGAGGAGGCTCTTCC 1410
 QY 1901 CTCCTCTCTAGAGAGAGGAGGAGTATTAACAACCTAGGGGGTGGGTAGGATTC 1960
 DB |||||||
 DB 1411 CTCCTCTCTAGAGAGGAGGAGGAGTATTAACAACCTAGGGGGTGGGTAGGATTC 1470

QY 1961 TAGGTATGGGAGAGATTGG 1982
Db 1471 TAGGTATGGGAGAGATTGG 1492

RESULT 6

ABL81903
ID ABL81903 standard; cDNA; 527 BP.

XX ABL81903;

DT 18-JUN-2002 (first entry)

XX Human colon cancer-associated cDNA, SEQ ID NO 10.

DE Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.

XX Homo sapiens.

XX WO200212280-A2.

PN 14-FEB-2002.

XX 30-JUL-2001; 2001WO-US23826.

XX 03-AUG-2000; 2000US-223265P.

PR 02-OCT-2000; 2000US-237406P.

PR 20-MAR-2001; 2001US-277495P.

PR 03-JUL-2001; 2001US-302702P.

XX (CORI-) CORIXA CORP.

PA Pyle RA, Xu J, Secrist H;

PI WPI; 2002-257462/30.

DR Novel polynucleotide encoding colon tumour polypeptides, useful as

PT vaccines for treating colon cancers

XX Claim 1; Page 147; 425pp; English.

XX The invention relates to isolated polynucleotides (I) encoding colon
CC tumour polypeptides (II). (I) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (I) are useful for determining the presence
CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (II) is useful for inhibiting
CC development of cancer in a patient. (I) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (I). ABL81903-ABL81904 represent human colon cancer cDNA
CC sequences of the invention.

XX SQ Sequence 527 BP; 86 A; 190 C; 158 G; 88 T; 5 other;

Query Match 22.4%; Score 444.6; DB 24; Length 527;
Best Local Similarity 88.1%; Pred. No. 4.9e-93;
Matches 518; Conservative 0; Mismatches 9; Indels 61; Gaps 1;

QY 47 GGGAGCCCTGGAGCCCGGCTGGCCCTCCCGCCCTGGGGTGCACATCGGCCCTGAGT 106

Db 1 GGGAGCNCCTGGAGCCCGGCTGGCCCTCCCGCCCTGGGGTGCACATCGGCCCTGAGT 60

QY 107 CCGGTCCAGGCTGTGGGCTGGGACGCGCCGCGCCACCGTGGCCAGGACGTGGGCGCTC 166

Db 61 CCGGTCCAGGCTGTGGGCTGGGACGCGCCGCGCCACCGTGGCCAGGACGTGGGCGCTC 89

QY 167 CTGCTTCCTCCAGGCGCCCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTTCCT 226

Db 90 -----GCCGCTGGGCGGAGTGGCCCGGCTTCCT 119

QY 227 GCCTTGGGCCACCTCTGCCCGCCGCTGGCCCTGGGGGCTCTGGTGTGGGCTCTTGGG 286
Db 120 GCCTTGGGCCACCTCTGCCCGCCGCTGGCCCTGGGGGCTCTGGTGTGGGCTCTTGGG 179
QY 287 GCTCTGGCAGCATCGCAGCCCGGCGGCTCCCATATGCGTGGGAGACGACCTG 346
Db 180 GCTCTGGCAGCATCGCAGCCCGGCGGCTCCCATATGCGTGGGAGACGACCTG 239
QY 347 CAGGACACGAGAAAGAAATATAGCCCGCCAGCCGACATCTGCTGCTCCGCTGCC 406
Db 240 CAGGACACGAGAAAGAAATATAGCCCGCCAGCCGACATCTGCTGCTCCGCTGCC 299
QY 407 GCCAGCACCTATGCTCAGCTAAATGTAGCCGCTCCGGGACACAGTTTGTGCCACATG 466
Db 300 GCCAGCACCTATGCTCAGCTAAATGTAGCCGCTCCGGGACACAGTTTGTGCCACATG 359
QY 467 TGCCGAGAAATTCCTACACGAGCAGTGAAGTACTGACCATCTGCCAGCTGTGCCGCCC 526
Db 360 TGCCGAGAAATTCCTACACGAGCAGTGAAGTACTGACCATCTGCCAGCTGTGCCGCCC 419
QY 527 CTGTGACCCAGTGTGGCCCTCGAGGAGATTGCCCTGTCACAAAGCAACGAGACCCA 586
Db 420 CTGTGACCCAGTGTGGCCCTCGAGGAGATTGCCCTGTCACAAAGCAACGAGACCCA 479
QY 587 GTGCCGCTGCCAGCCGGGAAATGTTCTGTGCTGCTGGGCGCTCGAGTG 634
Db 480 GTGCCGCTGCCAGCCGGGAAATGTTCTGTGCTGCTGGGCGCTCGAGTG 527

RESULT 7

ABL81903

ID ABL81903 standard; cDNA; 481 BP.

XX ABL81903;

DT 17-MAY-2002 (first entry)

DE Human ovarian cancer related cDNA clone SEQ ID NO:4881.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200192581-A2.

PD 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US17756.

XX 26-MAY-2000; 2000US-207484P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.

XX Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide

PS Claim 1; SEQ ID 4881; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
CC (S1) from the 10912 nucleotide sequences as given in ABU77023 to
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
CC population of (II), or antigen presenting cells that express (II).
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises

CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

XX Sequencé 481 BP; 113 A; 161 C; 123 G; 84 T; 0 other;

Query Match 21.9%; Score 433.6; DB 24; Length 481;
 Best Local Similarity 96.5%; Pred. No. 1.6e-90;
 Matches 465; Conservative 0; Mismatches 14; Indels 3; Gaps 2;

QY 390 TCGTCTCCGCTGCGCGCAGGACCTATGCTCAGCTAAATGTAGCGCATCCGGGAC 449

Db 1 TCGTCTCCGCTGCGCGCAGGACCTATGCTCAGCTAAATGTAGCGCATCCGGGAC 59

QY 450 ACAGTTTGTGCGACATGTGCGGAGATTCCTACAAGGACCTGGAACCTGACCATC 509

Db 60 ACAGTTTGTGCGACATGTGCGGAGATTCCTACAAGGACCTGGAACCTGACCATC 119

QY 510 TCCAGCTGTGCGCGCCTGTGACCCAGTGTGCGGCTCGAGGATTCGCCCTTCGACA 569

Db 120 TCCAGCTGTGCGCGCCTGTGACCCAGTGTGCGGCTCGAGGATTCGCCCTTCGACA 179

QY 570 AGCAACGGAGAACCCAGTGGCGTGCAGCGGGGAAATGTTCTGTGCTGGGCCCTC 629

Db 180 AGCAACGGAGAACCCAGTGGCGTGCAGCGGGGAAATGTTCTGTGCTGGGCCCTC 239

QY 630 GAGTGTACACACTGCGAGTACTTTCTGACTGCCCGCTGGCACTGAAGCGAGCTCAA 689

Db 240 GAGTGTACACACTGCGAGTACTTTCTGACTGCCCGCTGGCACTGAAGCGAGCTCAA 299

QY 690 GATGAAGTTGGAGGGTAAACACCACTGCGTCCCTGCAAGGCGAGGCACTT--CCAGA 747

Db 300 GATGAAGTTGGAGGGTAAACACCACTGCGTCCCTGCAAGGCGAGGCACTTTCAGAA 359

QY 748 ATACCTCTCCCGCAGCGCGCTGCCAGCCGCCACACAGGTGTGAGAACCAAGTCTGG 807

Db 360 GTACCTCTCCCGCAGCGCGCTGCCAGCCGCCACACAGGTGTGAGAACCAAGTCTGG 419

QY 808 TGGAGGCGAGTCCAGGCACTGCGGAGTCCGACACCACTGCAAAAATCCATTAGAGCCAC 867

Db 420 TGGAGGCGAGTCCAGGCACTGCGGAGTCCGACACCACTGCAAAAATCCATTAGAGCCAC 479

QY 868 TG 869

Db 480 TG 481

RESULT 8
 ABL82725

XX ABL82725 standard; cDNA; 529 BP.

AC ABL82725;

XX 17-MAY-2002 (first entry)

XX Human ovarian cancer related cDNA clone SEQ ID NO:5703.

XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US17756.

XX 26-MAY-2000; 2000US-207484P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.

XX Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide

PS Claim 1; SEQ ID 5703; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 CC (SI) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (SI) and/or (II) is useful for stimulating and/or expanding T cells
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

XX Sequence 529 BP; 84 A; 185 C; 149 G; 109 T; 2 other;

Query Match 19.5%; Score 386.4; DB 24; Length 529;
 Best Local Similarity 88.8%; Pred. No. 1.3e-79;
 Matches 474; Conservative 0; Mismatches 53; Indels 7; Gaps 5;

QY 76 CCGGCGCTGGGTCACATCGGCCTGAGTCCCTCCAGGCTCTGGGCTGGGCGAGCC 135

Db 1 CCGGCGCTGGGTCACATCGGCCTGAGTCCCTCCAGGCTCTGGGCTGGGCGAGCC 60

QY 136 GCGGCGACCGTGGCCAGGAGCTGGGCTCTCTGCTTCCCTCCAGGCGCCCGGCTGT 195

Db 61 G-CGCCANCGTGGCCAGGAGCTGGGCTCTCTGCTTCCCTCCGCTCTCTTTTTC 119

QY 196 GCGGCGCTGGCGAGTGGCCCGCCATG--CTCCTGCTTGGGCGACCTCTGCCCGGCT 253

Db 120 CCCTCTCGCGAGCTGGCGGCTCATGCTCTTGGGGGACCTGTGCTCCCGGCT 179

QY 254 GCGCTGGGCGCTCTGGTGTGGGCTCTCTGCGGCTCTCTGCGAGCATCGAGCCCGCAGC 313

Db 180 GCGCTGGGCG--CTCTGGTGTGGGCTCTCTGCGGCTCTCTGCGAGCATCGAGCCCGCAG-- 236

QY 314 GGTGCTCCATATGCTCGGAGAACACAGACCTGCGAGGACCGAGGAGGAAATGATGA 373

Db 237 CGGTGCTCCATATGCTCGGAGAACACAGACCTGCGAGGACCGAGGAGGAAATGATGA 296

QY 374 GCGGCGACCGCATCTGCTCTCCCGTGGCGCGGCGAGCACCTATGCTCAGCTAAATG 433

Db 297 GCGGCGACCGCATCTGCTCTCCCGTGGCGCGGCGAGCACCTATGCTCAGCTAAATG 355

QY 434 TAGCGCATCGGGACACAGTTTGTGCCACATGTGCCAGAAATTCCTACAACGAGCACTG 493

Db 356 TAGCGCATCGGGACACAGTTTGTGCCACATGTGCCAGAAATTCCTACAACGAGCACTG 415

QY 494 GAATCTACTGACCATCTGCAGCTGTGCCGCCCTGTGTGACCCAGTGTATGGCCCTCGAGGA 553
|||||
Db 416 GAATCTACTGACCATCTGCAGCTGTGCCGCCCTGTGTGACCCAGTGTATGGCCCTCGAGGA 475
QY 554 GATTGCCCTGTGACAGCAAGCAAGCGGAGACCCAGTGTGCCGTGCCAGCCGGGAAT 607
|||||
Db 476 GATTGCCCTGTGACAGCAAGCAAGCGGAGACCCAGTGTGCCGTGCCAGCCGGGAAT 529

RESULT 9
ABL82484
ID ABL82484 standard; cDNA; 516 BP.
XX
AC
ABL82484;
XX
17-MAY-2002 (first entry)
DE Human ovarian cancer related cDNA clone SEQ ID NO:5462.
XX
DE Human ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200192581-A2.
PN
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US17756.
PF
XX
XX 26-MAY-2000; 2000US-207484P.
PR
XX
XX (CORI-) CORIXA CORP.
PA
XX
PI Algate PA, Harlocker SL, Jones R;
XX
XX WPI; 2002-122075/16.
DR
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising polypeptide of a ovarian tumor polypeptide, polynucleotide encoding polypeptide, antibody specific to polypeptide or T cell expressing polypeptide -
XX
XX Claim 1; SEQ ID 5462; 489pp; English.
PS
XX
XX The present invention describes a composition (I) comprising: carriers and immunostimulants; and a polypeptide (II) of a ovarian tumour polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1) from the 10912 nucleotide sequences as given in ABL77023 to ABL87934, (III) encoding (II) having a sequence (S2), a T cell population of (II), or antigen presenting cells that express (II). (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to (S1) can be used for detecting ovarian cancer in a patient's biological sample preferably serum or ovarian tissue. The method comprises contacting a biological sample from a patient with (IV), detecting the amount of polynucleotide hybridising to (IV) and comparing the amount to a predetermined cutoff value and thereby detecting ovarian cancer in the patient, where the amount of polynucleotide hybridising to (IV) is detected preferably by polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is useful for stimulating and/or expanding T cells specific for an ovarian tumour protein comprising contacting T cells with (III) or (II). (III) is useful in design and preparation of ribozyme molecules for inhibiting expression of the tumour polypeptides and proteins in tumour cells; and to isolate a full length gene from a suitable library e.g., a tumour cDNA library using well known techniques.
XX
XX Sequence 516 BP; 83 A; 187 C; 153 G; 93 T; 0 other;

Query Match 19.3%; Score 382; DB 24; Length 516;
Best Local Similarity 90.5%; Pred. No. 1.3e-78;
Matches 476; Conservative 0; Mismatches 40; Indels 10; Gaps 6;

QY 95 TCGGCCCTGAGTCCCGTCCAGGCTCTGGGCTCGGCAGCGCCGCCACCGCTGCCCAGG 154
|||||
Db 1 TCGGCCCTGAGTCCCGTCCAGGCTCT-GGCTCGGGCAGCGG-CGCCACCGCTGCCCAGG 58
QY 155 ACGTGCGGCCCTCTGCTTCCCTCCAGGCCGCCACGTTGCTGGCCGCTGCCGAGTGGC 214
|||||
Db 59 ACGTGCGGCCCTCTGCTTCCCTCCAGGCCGCCACGTTGCTGGCCGCTGCCGAGGTGGC 118
QY 215 CGCATCTCTCTGCTTGGGCCACCTCTGCCGCCGCTGGCTTGGGGCTCTGTGGTCT 274
|||||
Db 119 GCGATGCTCTGCTTGGGCCACCTCTGCCGCCGCTGGCTTGGGGCTCTGTGGTCT 175
QY 275 GGGCTCTTTCGGGCTCTGTCAGCATCGCAGCCGCCAGCGGTGCTTCATATGCGTCGGA 334
|||||
Db 176 GGGCTCTTTCGGGCTCTGTCAGCATCGCAGCCGCCAGCGGTGCTTCATATGCGTCGGA 235
QY 335 GAACCCAGCTGTCAGGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 394
|||||
Db 236 GAACCCAGCTGTCAGGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 293
QY 395 CTCCGCTGCGGCCAGGACCTATGCTCAGCTAAATGTAGCGCATCCGGGACACAGT 454
|||||
Db 294 CGCCGCTGCGGCCAGGACCTATGCTCAGCTAAATGTAGCGCATCCGGGACACAGT 350
QY 455 TTGTGCCACATGTGCCGAGAAATTCCTACACGAGCACTGGAACCTACCTGACCATCTGCCA 514.
Db 351 TTGTGCCACATGTGCCGAGAAATTCCTACACGAGCACTGGAACCTACCTGACCATCTGCCA 410
QY 515 GCTGTGCCGCCCTGTGACCCAGTGTGATGGGCTCGAGGAGATGCCCCCTGCAACGAA 574
|||||
Db 411 GCTGTGCCGCCCTGTGACCCAGTGTGATGGGCTCGAGTGGCTCGATGAGATTGCCCTGCAACGAA 470
QY 575 ACGGAGACCCAGTGTGCCGCTGCCAGCGGGAATGTTCTGTGCTGCC 620
|||||
Db 471 ACGGAGACCCAGTGTGCCGCTGCCAGCGGGAATGTTCTGTGCTGCC 516

RESULT 10
ABL82623
ID ABL82623 standard; cDNA; 281 BP.
XX
AC ABL82623;
XX
XX 17-MAY-2002 (first entry)
DT
DE Human ovarian cancer related cDNA clone SEQ ID NO:5601.
XX
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX Homo sapiens.
XX WO200192581-A2.
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US17756.
XX
XX 26-MAY-2000; 2000US-207484P.
XX
XX (CORI-) CORIXA CORP.
PA
XX Algate PA, Harlocker SL, Jones R;
PI
XX WPI; 2002-122075/16.
DR
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising polypeptide of a ovarian tumor polypeptide, polynucleotide encoding polypeptide, antibody specific to polypeptide or T cell expressing polypeptide -
XX
XX Claim 1; SEQ ID 5601; 489pp; English.
PS
XX
XX The present invention describes a composition (I) comprising: carriers

CC and immunostimulants; and a polypeptide (II) of an ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 CC (SI) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

XX SQ Sequence 281 BP; 58 A; 94 C; 77 G; 52 T; 0 other;

Query Match 12.8%; Score 253.8; DB 24; Length 281;
 Best Local Similarity 97.1%; Pred. No. 4.4e-49;
 Matches 269; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 996 CCATTCTCGAGATGTTCCCGAGTATCCACTGGGCTCCCGCAGCCC-CAGTTTGGGA 1054
 || |||||
 DB 5 CCATTCTCGAGATGTTCCCGAGTATCCACTGGGCTCCCGCAGCCCAGTTTGGGA 64
 || |||||

QY 1055 GCGAGGGTGCCGCAACAGCAGAGTCTCTGGACCTGACCGAGGAGCGCGAGTTGGAACC 1114
 || |||||
 DB 65 GCGAGGGTGCCGCAACAGCAGAGTCTCTGGACCTGACCGAGGAGCGCGAGTTGGAACC 124
 || |||||

QY 1115 CGGGAGCAGACCGAGTGGCCCGGACGATCCCAATGGCATTCATGTCACCGCGGGTCTAT 1174
 || |||||
 DB 125 CGGGAGCAGACTCAGTGGGCCCGACGCTACCAATGGCATTCATGTCACCGCGGGTCTAT 184
 || |||||

QY 1175 GACTATCACTGGCAACATCTACATCTACAATGGACAGTACTGGGGGACACCGGGTCC 1234
 || |||||
 DB 185 GACTATCACTGGCAACATCTACATCTACAATGGACAGTACTGGGGGACACCGGGTCA 244
 || |||||

QY 1235 TGGAGACTCCCGAGTACCCCGCAACCTCCATACCCC 1271
 || |||||
 DB 245 TGGAGACTCCCGAGTACCCCGCAACCTCCATTCGCC 281
 || |||||

RESULT 11
 ABL82688/c
 ID ABL82688 standard; cDNA; 289 BP.
 XX AC ABL82688;
 XX DT 17-MAY-2002 (first entry)
 XX DE Human ovarian cancer related cDNA clone SEQ ID NO:5666.
 XX KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200192581-A2.
 XX PD 06-DEC-2001.
 XX PF 29-MAY-2001; 2001WO-US17756.
 XX PR 26-MAY-2000; 2000US-207484P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.
 XX Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of an ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, antibody specific to polypeptide or T cell expressing
 PT polypeptide
 XX Claim 1; SEQ ID 5666; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of an ovarian tumour
 CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
 CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 CC (SI) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (III) or (II). (III) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.

XX SQ Sequence 289 BP; 49 A; 75 C; 93 G; 72 T; 0 other;

Query Match 12.7%; Score 251.2; DB 24; Length 289;
 Best Local Similarity 98.5%; Pred. No. 1.7e-48;
 Matches 264; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1007 AGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGAGTTTGGAGGAGGGTGCC 1066
 || |||||
 DB 289 AGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGAGTTTGGAGGAGGGTGCC 230
 || |||||

QY 1067 GCAACAGCAGA-GTCTCTGGACCTGACAGGAGCGCGAGTGTGAACCCGGGAGCAGA 1125
 || |||||
 DB 229 GCAACAGCAGATGTCTCTGGACCTGACAGGAGCGCGGTGTGAACCCGGGAGCAGA 170
 || |||||

QY 1126 GCCAGTGGCCCGAGGTACCAATGGCATTCATGTACCGGCGGGTCTATGACTATCACTG 1185
 || |||||
 DB 169 GCCAGTGGCCCGAGGTACCAATGGCATTCATGTACCGGCGGGTCTATGACTATCACTG 110
 || |||||

QY 1186 GCAACATCTACATCTACAATGGACAGTACTGGGGGACACCGGGTCTCTGGAGACCTCC 1245
 || |||||
 DB 109 GCAACATCTACATCTACAATGGACAGTACTGGGGGACACCGGGTCTCTGGAGACCTCC 50
 || |||||

QY 1246 CAGTACCCCGGAACCTCCATFACCCCAT 1273
 || |||||
 DB 49 CAGTACCCCGGAACCTCCATFACCCCAT 22
 || |||||

RESULT 12
 AAQ45225
 ID AAQ45225 standard; cDNA; 1557 BP.
 XX AC AAQ45225;
 XX DT 07-OCT-1994 (first entry)
 XX DE Sequence encoding a recombinant human tumour necrosis factor receptor
 DE TNFR/fc fusion protein.
 XX KW Tumour necrosis factor receptor; chimeric antibody molecule;
 XX KW immunoglobulin; ss.

OS Synthetic.
 XX Key Location/Qualifiers
 FH mat_peptide 1..1554
 FT /*tag= a
 FT
 XX
 XX WO9406476-A.
 XX
 XX 31-MAR-1994.
 XX
 XX 14-SEP-1993; 93WO-US08666.
 XX
 XX 15-SEP-1992; 92US-0946236.
 XX
 XX (IMMUNEX CORP.
 XX
 XX Jacobs CA, Smith CA;
 XX
 XX WPI; 1994-118172/14.
 XX P-PSDB; AAR51003.
 XX
 XX Treating TNF mediated inflammatory diseases with TNF antagonist -
 PT esp. soluble form of TNF receptor, opt. as fusion protein with
 PT human immunoglobulin Fc region, esp. for treating arthritis
 XX
 XX Disclosure; Page 32-34; 47pp; English.
 XX
 XX AAQ45224 is cDNA from clone 1 of library WI-26 V44 of human fibroblast
 CC cell line WI-26 V44. The mature full-length TNFR1 is a glycoprotein
 CC having a mol. wt. of about 75-80 kDa. The cloning of the cDNA for
 CC TNFR1 was described in Smith et al., Science 248:1019, 1990. Clone 1
 CC is contained in expression vector pCAV/NOT-TNFR (ATCC 68088). A
 CC recombinant chimeric antibody may be produced having TNFR sequences
 CC substituted for the variable domains of either or both of the
 CC immunoglobulin molecule heavy and light chains and having
 CC unmodified constant region domains. A specific example of a TNFR/Fc
 CC fusion protein is given in AAQ45225/R51003. The rhu TNFR:Fc fusion
 CC gene was created by ligating the following fragments into a cloning
 CC vector: 1) an 867 bp Asp718-Pvu2 fragment from pCAV/NOT-TNFR (ATCC
 CC 68088) contg. the cDNA encoding the truncated TNFR, 2) a 700 bp
 CC StyI-SpeI fragment from plasmid pIXY498 coding for 232 AAs of the
 CC Fc portion of human IgG1. Plasmid pIXY498 is a yeast expression
 CC vector contg. the Fc fragment of human IgG1. 3) An oligo linker, to
 CC fuse the truncated TNFR with the human IgG1 Fc fragment. This
 CC linker was created by PCR using primer AAQ45226, which encodes the 3'
 CC end of the truncated TNF receptor and the 5' end of human IgG1 and
 CC primer AAQ45227, which is an antisense sequence encoding bps 257-237
 CC of human IgG1.
 XX
 XX Sequence 1557 BP; 345 A; 524 C; 445 G; 243 T; 0 other;
 SQ
 Query Match 3.6%; Score 72; DB 15; Length 1557;
 Best Local Similarity 51.3%; Pred. No. 5.7e-07;
 Matches 193; Conservative 0; Mismatches 180; Indels 3; Gaps 1;
 QY 241 CTGCCCCCGCGCTGGCGCTCCATATGCGTCGGAGACACAGCTGCAGGACAGCAT 300
 DB 95 CCGTCGCGCTGCTGGCGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCT 154
 QY 301 CGCAGCCCCAGGCGGCTGCCATATGCGTCGGAGACACAGCTGCAGGACAGGAAA 360
 DB 155 TGCCCGCCAGGTGGGATTTACACCTTACGCCCGGAGCCCGGAGACATGCCGGTCA 214
 QY 361 AGGAATACTATGAGCCCCAGCAGCATCTGCTGCTCCCGCTGCCCGGACCATATG 420
 DB 215 GAGAACTACTATGACAGACAGCTCAGATGTGCTGCAGCAATGCTGCCGGGCGACATG 274
 QY 421 TCTCAGCTAAATAGCCGATCCGGGACACAGTTTGTGCCACATGTGCCAGAAATTCCT 480
 DB 275 CAAAAGTCTTCTGACCAAGACCTCGGACACCTGTGTGACTCTGTGAGGACAGCAT 334
 QY 481 ACAACGAGCACTGGAACCTACCTGACCATGTGCCAGCTGTGCCCGCCCTGTGACCCAGTGA 540
 DB 361 AGGAATACTATGAGCCCCAGCAGCATCTGCTGCTCCCGCTGCCCGGACCATATG 420

335 ACACCCAGCTCTGGAACCTGGGTTCCCGAGTGTAGCTGTGGCTCCCG---CTGTAGCT 391
 541 TGGGCTCGAGGAGATGCCCCCTGACACAGCAACAGGAGACCCAGTGGCGTGCAGC 600
 392 GTGACCAGGTGGAACTCAAGCTGCACCTCGGGAACAGAACCCGATCTGCACCTGCAGC 451
 601 CGGGAATGTTCTGTGC 616
 452 CCGGCTGTTACTGGC 467

RESULT 13
 AAC90446
 ID AAC90446 standard; cDNA; 1557 BP.
 XX
 XX AAC90446;
 XX
 XX 19-MAR-2001 (first entry)
 XX
 XX TNFR:Fc fusion protein coding sequence.
 XX
 XX Antiinflammatory; antiarthritic; gene therapy; inflammatory disorder;
 KW Tumour Necrosis Factor receptor; TNFR; IgG1 Fc domain; asthma;
 KW rheumatoid arthritis; Crohn's disease; congestive heart failure; ss.
 XX
 XX Rattus sp.
 OS
 XX WO200073481-A1.
 PN
 XX 07-DEC-2000.
 PD
 XX 26-MAY-2000; 2000WO-US14586.
 PF
 XX 28-MAY-1999; 99US-0150688.
 PR
 XX (TARG-) TARGETED GENETICS CORP.
 PA
 XX Burstein H, Stepan AM;
 PI
 XX WPI; 2001-061552/07.
 DR
 XX P-PSDB; AAB50080.
 XX
 XX Recombinant adeno-associated virus vector, useful for palliating tumor
 PT necrosis factor associated disorder, comprises polynucleotide which
 PT encodes fusion polypeptide comprising tumor necrosis factor receptor
 XX
 XX Example 1; Fig 2; 90pp; English.
 XX
 XX The present invention relates to a recombinant adeno-associated virus
 CC (rAAV) vector comprising a polynucleotide (the present sequence) which
 CC encodes a fusion protein comprising an extracellular domain of Tumor
 CC Necrosis Factor receptor (TNFR) and a constant domain (Fc) of IgG1
 CC molecule. The vector of the present invention is useful for reducing TNF
 CC levels, for reducing an inflammatory response which occurs in a
 CC connective tissue and for palliating a TNF-associated disorder such as
 CC inflammatory disorders e.g. rheumatoid arthritis, Crohn's disease, asthma
 CC and congestive heart failure.
 XX
 XX Sequence 1557 BP; 345 A; 526 C; 445 G; 241 T; 0 other;
 SQ
 Query Match 3.6%; Score 72; DB 22; Length 1557;
 Best Local Similarity 51.3%; Pred. No. 5.7e-07;
 Matches 193; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

```

|||||
Db 215 GAGAACTATGACGACAGACAGCTCAGATGTGCTGCAGCAATGCTCGCGGCGCAACATG 274
QY 421 TCTCAGCTAAATGTAGCGGATCGGACACAGATTTGTGCCACATGTGCCGAGAAATTCCT 480
Db 275 CAAAGTCTTCTGTACCAAGACCTCGGACACCGGTGTGACTCTGTGAGGACAGACAT 334
QY 481 ACAACGAGCACTGGAATACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTGA 540
Db 335 ACACCCAGCTCTGAACCTGGTTCGCCAGTGTGAGCTGTGGCTCCG---CTGTAGCT 391
QY 541 TGGCCCTCGAGGAGATTGCCCTTGCACAAAGCAACGGAACCCAGTGTGCCGCTGCCAGC 600
Db 392 CTGACCCAGTGGAACTCAAGCCTGCATCGGGAACAGAACCCGATCTGCACCTGCAGGC 451
QY 601 CGGGAATGTTCTGTGC 616
Db 452 CCGCTGTGTACTGCGC 467

RESULT 14
AAQ10990
ID AAQ10990 standard; cdna to mRNA; 1640 BP.
XX AC AAQ10990;
XX DT 24-MAY-1991 (first entry)
XX DE Partial sequence of hTNF-R clone 1.
XX KW Tumour necrosis factor receptor; immune response; inflammation;
XX KW cachexia; septic shock; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 88..1473
XX FT mat_peptide 154..1470
XX FT sig_peptide 88..153
XX FT /*tag= c
XX EP418014-A.
XX PN 20-MAR-1991.
XX PD 10-SEP-1990; 90EP-0309875.
XX PF 10-MAY-1990; 90US-0523635.
XX PR 11-SEP-1989; 89US-0405370.
XX PR 13-OCT-1989; 89US-0421417.
XX XX (IMMU-) IMMUNEX CORP.
XX PI Smith CA, Goodwin RG, Beckmann PM;
XX WPI: 1991-082230/12.
XX P-PSDB: AAR11141.
XX New tumour necrosis factor -alpha and -beta receptors - and DNA
XX encoding these used to regulate immune responses in treatment of
XX cachexia, septic shock or side-effects of cytokine therapy.
XX Claim 1; Fig 2; 41pp; English.
XX The sequence was obtd. from a clone isolated from library prepd.
XX from a human fibroblast cell line, WI-26 VA4 (ATCC CCL 95.1).
XX The clone is deposited as Accession No. 68088 under the name
XX pCAV/NOT-TNF-R. The DNA can be truncated to produce sequences which
XX express soluble receptor comprising residues 1-235, 1-185 or 1-163
XX of the protein.
XX See also AAQ10991.

```

```

XX SQ Sequence 1640 BP; 328 A; 543 C; 502 G; 267 T; 0 other;
Query Match 3.6%; Score 72; DB 12; Length 1640;
Best Local Similarity 51.3%; Pred. No. 5.8e-07;
Matches 193; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 241 CTGCCCCCGCGCTGGCCCTCTGGTGTGGGCTCTTTCGGGCTCCTTGGCAGCAT 300
Db 95 CCGTCGCGCTCTGGCCCGCGCTGGCCCTCGGACTTGGGCTCGGCGCACGCT 154
QY 301 CGCAGCCCCAGGCGGTGCCCTCATATGCTCGGAGAACACAGACCTCGAGGAGACAGGAAA 360
Db 155 TGCCTCGCCAGGTGGCATTTACACCTTACGCCCGGAGCGCGGAGACACATGCCGGCTCA 214
QY 361 AGGAATACTATGAGCCCCAGCAGCATCTGCTCCCGCTGCCCGCAGGCACTATG 420
Db 215 GAGAACTATGACGACAGCTCAGATGTGCTGCAGCAATGCTCGCGGCGCAACATG 274
QY 421 TCTCAGCTAAATGTAGCGGATCGGACACAGATTTGTGCCACATGTGCCGAGAAATTCCT 480
Db 275 CAAAGTCTTCTGTACCAAGACCTCGGACACCGGTGTGACTCTGTGAGGACAGACAT 334
QY 481 ACAACGAGCACTGGAATACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTGA 540
Db 335 ACACCCAGCTCTGGAACCTGGTTCGCCAGTGTGAGCTGTGGCTCCG---CTGTAGCT 391
QY 541 TGGGCTCGAGGAGATTGCCCTTGCACAAAGCAACGGAACCCAGTGTGCCGCTGCCAGC 600
Db 392 CTGACCCAGTGGAACTCAAGCCTGCATCGGGAACAGAACCCGATCTGCACCTGCAGGC 451
QY 601 CGGGAATGTTCTGTGC 616
Db 452 CCGCTGTGTACTGCGC 467

RESULT 15
AAQ49931
ID AAQ49931 standard; cdna to mRNA; 1641 BP.
XX AC AAQ49931;
XX DT 29-APR-1994 (first entry)
XX DE TNF-R cdna.
XX KW Human; tumour necrosis factor receptor; TNF-R; interleukin-1 receptor;
XX KW IL-1R; fusion protein; linker; TNF; IL-1; cachexia; cerebral malaria;
XX KW rheumatoid arthritis; diabetes; multiple sclerosis; septic shock;
XX KW pulmonary fibrosis; silicosis; allograft; xenograft; rejection;
XX KW graft versus host disease; sepsis; inflammation; allergy;
XX KW autoimmune dysfunction; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 88..1473
XX FT /*tag= a
XX FT /product= hTNF-R
XX FT sig_peptide 88..153
XX FT mat_peptide 154..1470
XX FT /*tag= b
XX FT /*tag= c
XX PN W09319777-A.
XX PD 14-OCT-1993.
XX PF 26-MAR-1993; 93WO-US02938.
XX PR 30-MAR-1992; 92US-0860710.
XX XX (IMMU-) IMMUNEX CORP.

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:09:45 ; Search time 2289.32 Seconds
(without alignments)
14021.414 Million cell updates/sec

Title: US-09-917-372-2
Perfect score: 1982
Sequence: 1 gccccgcgcagctcgtc.....ggtatggggaagaggtttgg 1982

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpi:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	861.2	43.5	954	14	BQ930450
2	821.6	41.5	882	14	BQ691803
3	818	41.3	933	9	AL564542
4	800	40.4	1006	14	BM923844
5	798.2	40.3	1109	13	BM463063
6	797.4	40.2	978	9	AL564472

7	794.8	40.1	993	9	AL513838
8	794.2	40.1	816	13	BI908262
9	785.6	39.6	867	13	BI768812
10	755.4	38.1	840	13	BI906190
11	746.6	37.7	954	13	BI821009
12	729	36.8	895	14	BQ957322
13	723.2	36.5	891	13	BI911211
14	716.8	36.2	788	13	BI818683
15	714.4	36.0	881	13	BI761249
16	713.4	35.0	808	13	BI818411
17	710.4	35.8	916	13	BI908596
18	708.4	35.7	968	14	BQ653019
19	704	35.5	917	13	BI760125
20	698.6	35.2	932	13	BI765370
21	696.4	35.1	801	12	BF792867
22	695.2	35.1	1102	14	BM925105
23	690.8	34.9	710	13	BI907277
24	681.6	34.4	933	14	BQ214865
25	679.8	34.3	801	9	AI193089
26	679.6	34.3	918	13	BI770354
27	676.6	34.1	764	12	BG548459
28	670.2	33.8	749	12	BF789430
29	668.4	33.7	735	12	BG331666
30	665.8	33.6	942	13	BI818552
31	665	33.6	1239	14	BM923679
32	664.4	33.5	666	14	BM767318
33	660	33.3	732	9	AUI38624
34	649.6	32.8	794	9	AI193092
35	638.4	32.2	780	9	AUI36112
36	637.8	32.2	814	9	AUI18203
37	636	32.1	636	14	BM849181
38	633	31.9	742	14	BQ000587
39	632.6	31.9	968	14	BQ898015
40	632	31.9	681	9	AUI39374
41	630.4	31.8	920	12	BG026287
42	629.2	31.7	634	14	BM746911
43	626	31.6	841	12	BG750836
44	622	31.4	622	14	BQ636409
45	620.4	31.3	721	14	BM985373

ALIGNMENTS

RESULT 1 BQ930450 954 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT_8946139 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6462494
DEFINITION 5', mRNA sequence.
ACCESSION BQ930450
VERSION BQ930450.1 GI:22345481
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2652 row: h column: 15
High quality sequence stop: 576.
Location/Qualifiers
1. .954

FEATURES
source

QY	1065	CCGCAACAGCAGAGTCTCTGTGACCTGACCAAGGAGCGCGAGTTGGAACCCGGGGAGCAG	1124
Db	1	CCGCAACAGCAGAGTCTCTGTGACCTGACCAAGGAGCGCGAGTTGGAACCCGGGGAGCAG	60
QY	1125	AGCCAGTGGCCACGAGTACCAATGCAATTCATCTACCGCGGGTCTATGACATATCACT	1184
Db	61	AGCCAGTGGCCACGAGTACCAATGCAATTCATCTACCGCGGGTCTATGACATATCACT	120
QY	1185	GGCAACATCTACATCTACAATGGACCAAGTACTGGGGGGACCCACCGGGTCTCTGGAGACCTC	1244

```
Db 121 GGCACATCTACATCTACATGACAGTACTGGGGGACACACCGGCTCTCGGAGACCTC 180
Qy 1245 CCAGCTACCCCGGAACTCCATACCCATCCCGAAGAGGGGAGCCCTGGCCCTCCCGGG 1304
Db 181 CCAGTACCCCGGAACTCCATACCCATCCCGAAGAGGGGAGCCCTGGCCCTCCCGGG 240
Qy 1305 CTCTCTACACCCACAGGAGATGGCAAGCTTGGACCTAGCGGAGACAGACACTGT 1364
Db 241 CTCTCTACACCCACAGGAGATGGCAAGCTTGGACCTAGCGGAGACAGACACTGT 300
Qy 1365 GGTGCCACACCTCTAACAGGGGCCCAAGGAACCAATTTATCACCCATGACTGACTGAGT 1424
Db 301 GGTGCCACACCTCTAACAGGGGCCCAAGGAACCAATTTATCACCCATGACTGACTGAGT 360
Qy 1425 CTGAGAAAGCGCAGAGAGGGGGGCAAGAGGACACTTCTCCCTTGAGGCTGCCCTGCC 1484
Db 361 CTGAGAAAGCGCAGAGAGGGGGGCAAGAGGACACTTCTCCCTTGAGGCTGCCCTGCC 420
Qy 1485 CACGTGGGATTCACAGGGGCTGAGTAGGGCCCGGGGAGCAGAGCCCTAAGGGATTAA 1544
Db 421 CACGTGGGATTCACAGGGGCTGAGTAGGGCCCGGGGAGCAGAGCCCTAAGGGATTAA 480
Qy 1545 GCTCAGACACCTCTGAGAGCAGGTGGGCACCTGGGTACGGTGCCTCCACAGGACTC 1604
Db 481 GCTCAGACACCTCTGAGAGCAGGTGGGCACCTGGGTACGGTGCCTCCACAGGACTC 540
Qy 1605 TCCCTACTGCTGAGCAAACTGAGGCTCCCGGCAGACCCACCCCTGGGGGCTGC 1664
Db 541 TCCCTACTGCTGAGCAAACTGAGGCTCCCGGCAGACCCACCCCTGGGGGCTGC 600
Qy 1665 TCAGCCTCAGGCAGCGGACAGGACATGATACCACTGTGCTGCCACTACGSCACGCGCA 1724
Db 601 TCAGCCTCAGGCAGCGGACAGGACATGATACCACTGTGCTGCCACTACGSCACGCGCA 660
Qy 1725 CCGGAGCAGCGCAGGAGCGCCGACAGCGTCACCTGCAAGGAGTCAAGGCGCCCT 1784
Db 661 CCGGAGCAGCGCAGGAGCGCCGACAGCGTCACCTGCAAGGAGTCAAGGCGCCCT 720
Qy 1785 CTAAAGGATTGCT-GGTGCTCATCCCAAGCTTCAGAGACCTTTGGGGTTCCACACTTC 1843
Db 721 CTAAAGGATTGCTGGGTGCTCATCCCAAGCTTCAGAGACCTTTGGGGTTCCACACTTC 780
Qy 1844 ACCTGGAGTGGTGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1899
Db 781 ACCTGGAGTGGTGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Qy 1900 CCTCCCTCTCTAGAGGAGGAGGAGGAGTCAATTAACA 1938
Db 841 TCCCTCTCTANNAGGAGAGGAGGAGTCAATTAACA 879

RESULT 3
AL564542 933 bp mRNA linear EST 16-FEB-2001
LOCUS AL564542 LTI_NFL001_NBC4 Homo sapiens cdna clone CS0DM007YN10 5
DEFINITION prime, mRNA sequence.
ACCESSION AL564542
VERSION AL564542.1 GI:12915052
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source location/Qualifiers
1. .933
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DM007YN10"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 163 a 335 c 257 g 166 t 12 others
ORIGIN

Query Match 41.3%; Score 818; DB 9; Length 933;
Best Local Similarity 97.9%; Pred. No. 6.9e-175;
Matches 826; Conservative 11; Mismatches 6; Indels 1; Gaps 1;

Qy 57 GAGCCCGCGCTGCGCGCTCCCGCCCTGGGGTGCACATCGGCCCTGAGTCCCTGCCAG 116
Db 1 GAGCCCGCGCTGCGCGCTCCCGCCCTGGGGTGCACATCGGCCCTGAGTCCCTGCCAG 60
Qy 117 GCTCTGGGCTCGGCAGCCGCCCGCCCTGGCCAGGACGTGCGGGCTCTGCTTCTTCT 176
Db 61 GCTCTGGGCTCGGCAGCCGCCCGCCCTGGCCAGGACGTGCGGGCTCTGCTTCTTCT 120
Qy 177 CCCAGGCCCGCCACCTTGTGCGCGCTGGCGAGTGGCGGCATGCTTCCCTTGGGCC 236
Db 121 CCCAGGCCCGCCACCTTGTGCGCGCTGGCGAGTGGCGGCATGCTTCCCTTGGGCC 180
Qy 237 ACCTCTCCCGCCCGCTGCGCTGGGGCCCTGCTGCTGGGCCCTTTCGGGCTCTTGSCA 296
Db 181 ACCTCTCCCGCCCGCTGCGCTGGGGCCCTGCTGCTGGGCCCTTTCGGGCTCTTGSCA 240
Qy 297 GCATCGCAGCCCGCAGGCGGTGCTCCATATGCTGCGGAGAACAGACCTGCGAGGACGAG 356
Db 241 GCATCGCAGCCCGCAGGCGGTGCTCCATATGCTGCGGAGAACAGACCTGCGAGGACGAG 300
Qy 357 GAAAGGAATATATGAGCCCGCAGCAGCGCATCTGCTCCCGCTGCGCCCGCAGGACCC 416
Db 301 GAAAGGAATATATGAGCCCGCAGCAGCGCATCTGCTCCCGCTGCGCCCGCAGGACCC 360
Qy 417 TATGCTCAGCTAAATGAGCCGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAA 476
Db 361 TATGCTCAGCTAAATGAGCCGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAA 420
Qy 477 TCCTACACGAGCAGCTGGAACCTACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCA 536
Db 421 WCMTAAMAGCAGCTGGAACCTACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCA 480
Qy 537 GTGATGGCCCTCGAGGAGATTGCCCTCGACAAAGCAAGCAAGCAAGCAAGCAAGCAAG 596
Db 481 GTGATGGCCCTCGAGGAGATTGCCCTCGACAAAGCAAGCAAGCAAGCAAGCAAGCAAG 540
Qy 597 CAGCGGGAATGTTCTGCTGCTGCTGGGCCCTCGAGTGTACACACTGCGAGCTACTTCT 656
Db 541 CAGCGGGAATGTTCTGCTGCTGCTGGGCCCTCGAGTGTACACACTGCGAGCTACTTCT 600
Qy 657 GACTGCCCGCCCTGCGCAGCTGAAGCGGAGCTCAAGATGAAGTTGGGAGGGTAAACACAC 716
Db 601 GACTGCCCGCCCTGCGCAGCTGAAGCGGAGCTCAAGATGAAGTTGGGAGGGTAAACACAC 660
Qy 717 TGGCTCCCTCGCAGGAGGAGGAGCTTCCAGAAATACCTTCCCGCCAGCGCCGCTGCCAG 776
Db 661 TGGCTCCCTCGCAGGAGGAGGAGCTTCCAGAAATACCTTCCCGCCAGCGCCGCTGCCAG 720
Qy 777 CCCACACAGGTTGTGAGAACCAAGGTTCTGCTGGAGCAGCTCCAGGCACTGCCCGAGTCC 836
```

```

Db 721 CCCACACAGGTGTGAGAACCAAGTCTGGTGGAGGCAAGTCCAGGCACTGCCAGTCC 780
QY 837 GACACAACTGCAAAAATCCA-TTAGAGGCACCTGCCCCAGAGATGTCAGGATCGCTGCT 895
Db 781 GACACAACTGCAAAAATCCATTTAGAGCCACTGCCCCAGAGATGTCAGGAGCAATGCT 840
QY 896 CAAG 899
Db 841 GATG 844

RESULT 4
BM923844
LOCUS
DEFINITION
  AGENCOURT_6709962 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5759970
  5', mRNA sequence.
ACCESSION
  BM923844
VERSION
  BM923844.1 GI:19374223
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 1006)
REFERENCE
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: Life Technologies, Inc.
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLAM12806 row: d column: 19
  High quality sequence stop: 688.
  Location/Qualifiers
    1..1006
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:5759970"
      /clone_lib="NIH_MGC_116"
      /lab_host="DH10B"
      /note="Organ: pooled colon, kidney, stomach; Vector:
      pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
      source anonymous pool of 3 colons, age 26 yo male, 49 yo
      female, 71 yo male colon; 46 yo male kidney, and pool of 2
      stomachs, 62 yo male and 70 yo female. Library is
      oligo-dr primed and directionally cloned (EcoRV site is
      destroyed upon cloning). Average insert size 1.4 kb,
      insert size range 1-3 kb. Library is normalized and
      enriched for full-length clones and was constructed by C.
      Gruber (Invitrogen). Research Genetics tracking code
      023. Note: this is a NIH_MGC Library."
BASE COUNT
  186 a 364 c 276 g 180 t
ORIGIN

Query Match
  40.4%; Score 800; DB 14; Length 1006;
Best Local Similarity
  97.5%; Pred. No. 8.5e-17;
Matches 823; Conservative 0; Mismatches 20; Indels 1; Gaps 1;

QY 57 GAGGCGCGCGCTGGCGCTCCGGCGCTGGGTGCACATCGCGCCCTGAGTCGCGCCAG 116
Db 1 GAGGCGCGCGCTGGCGCTCCGGCGCTGGGTGCACATCGCGCCCTGAGTCGCGCCAG 60
QY 117 GCTCTGGGCTCGGCGAGCGCGCCACCGCTGCCAGGACGTCGGGCGCTCTGCTCTCT 176
Db 61 GCTCTGGGCTCGGCGAGCGCGCGCCACCGCTGCCAGGACGTCGGGCGCTCTGCTCTCT 120
QY 177 CCCAGGCCCCACAGTGTGCTGGCGCGCTGGCGAGTGGCGCCATGCTCTGCTGGGCC 236

```

```

Db 121 CCAGGCCCCACAGTTGCTGGCGCGCTGGCGAGTGGCGCCATGCTCTCTGCTGGCGC 180
QY 237 ACCTCTGCGCCCGCGCTGCGCTGGGGCCCTCTGGTCTCTGGGCTCTTTCGGGCTCTTGGCA 296
Db 181 ACCTCTGCGCCCGCGCTGCGCTGGGGCCCTCTGGTCTCTGGGCTCTTTCGGGCTCTTGGCA 240
QY 297 GCATCGACGCCCGCGCGCTGCTCCATATGCGTCCGAGAACACAGACCTGCAGGGACCCAG 356
Db 241 GCATCGACGCCCGCGCGCTGCTCCATATGCGTCCGAGAACACAGACCTGCAGGGACCCAG 300
QY 357 GAAAGGAATACTATGAGCCCGCAGCACCGCATCTGCTGCTCCCGTGCCTGCCCGCAGGCACC 416
Db 301 GAAAGGAATACTATGAGCCCGCAGCACCGCATCTGCTGCTCCCGTGCCTGCCCGCAGGCACC 360
QY 417 TATGCTCAGCTAAATGTAGCCGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAAT 476
Db 361 TATGCTCAGCTAAATGTAGCCGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAAT 420
QY 477 TCCTACACGAGCAGCTGGAACTACCTGACCATCTGCGAGCTGTGCCGCCCTGTGACCCA 536
Db 421 TCCTACACGAGCAGCTGGAACTACCTGACCATCTGCGAGCTGTGCCGCCCTGTGACCCA 480
QY 537 GTGATGGGCTCGAGGAGATTGCCCGCTGCACAAGCAAAACGGAAGACCCAGTGCCTGTC 596
Db 481 GTGATGGGCTCGAGGAGATTGCCCGCTGCACAAGCAAAACGGAAGACCCAGTGCCTGTC 540
QY 597 CAGCCGGGAATCTCTGCTGCTGCTGGCGCTTGCAGTGTACACACTGCGAGCTACTTCT 656
Db 541 CAGCCGGGAATCTCTGCTGCTGCTGGCGCTTGCAGTGTACACACTGCGAGCTACTTCT 600
QY 657 GACTGCCCGCTGGCAGCTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTAAACACCCAC 716
Db 601 GACTGCCCGCTGGCAGCTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTAAACACCCAC 660
QY 717 TCGTCTCCCTGCAAGCAGGCGACTTCCAGATACCTCTCCCGCAGCCCGCTGCGCAG 776
Db 661 TCGTCTCCCTGCAAGCAGGCGACTTCCAGATACCTCTCCCGCAGCCCGCTGCGCAG 720
QY 777 CCCCACACAGGTGTGAG-AACCAAGTCTGGTGGAGGAGCTCCAGGACACTGCCAGTC 835
Db 721 CCCCACACAGGTGTGAGAACCAAGTCTGGTGGAGGAGCTCCAGGACACTGCCAGTC 780
QY 836 CGACACAACTGCAAAAATCCATTAGAGCCACTGCCCGCAGAGATGTCAGGATCGCTGCT 895
Db 781 CGACACAACTGCAAAAATCCATTAGAGCCACTGCCCGCAGAGATGTCAGGAAACCATGCT 840
QY 896 CAAG 899
Db 841 GATG 844

RESULT 5
BM463063
LOCUS
DEFINITION
  AGENCOURT_6447515 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5537750
  5', mRNA sequence.
ACCESSION
  BM463063
VERSION
  BM463063.1 GI:18512103
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 1109)
REFERENCE
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: ATCC/DCTD/DTF
  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L1AM12229 row: i column: 15
High quality sequence stop: 659.

FEATURES

location/Qualifiers
1..1109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5537750"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 211 a 394 c 309 g 195 t
ORIGIN
Query Match 40.3%; Score 798.2; DB 13; Length 1109;
Best Local Similarity 98.1%; Pred. No. 2.2e-170;
Matches 839; Conservative 0; Mismatches 13; Indels 3; Gaps 3;
QY 28 ACTTCTGAGCTCCGCGCATGGAGCCCTGGAGCCCGGCTGGCGCTCCCGGCTCCCGGCTGGG 87
DB 2 ACTTCTGAGCTCCGCGCATGGAGCCCTGGAGCCCGGCTGGCGCTCCCGGCTCCCGGCTGGG 61
QY 88 GTGCACATCGGCGCTGAGTCCGCTCCAGGCTCTGGGCTGGGCGAGCGCCGCGCACCGT 147
DB 62 GTGCACATCGGCGCTGAGTCCGCTCCAGGCTCTGGGCTGGGCGAGCGCCGCGCACCGT 121
QY 148 GCCAGGAGCTGGGCTCTCGCTCTCCCTCCAGGCCCCAGTGTGCTGGCGCGCTGGCC 207
DB 122 GCCAGGAGCTGGGCTCTCGCTCTCCCTCCAGGCCCCAGTGTGCTGGCGCGCTGGCC 181
QY 208 GAGTGGCGCCCATGCTCTCGCTTGGGCGACCTCTGCTCCCGCGGCTGGGCGGCTG 267
DB 182 GAGTGGCGCCCATGCTCTCGCTTGGGCGACCTCTGCTCCCGCGGCTGGGCGGCTG 241
QY 268 TGGTGTGGGCTCTTGGGCTCTGGGCGATCGGAGCCCGAGCGGCTGCTCCATATG 327
DB 242 TGGTGTGGGCTCTTGGGCTCTGGGCGATCGGAGCCCGAGCGGCTGCTCCATATG 301
QY 328 CGTGCGAGAACAGACCTCGAGGACCGAGAAAGAAATATAGTATGAGCCCGAGCACCGCA 387
DB 302 CGTGCGAGAACAGACCTCGAGGACCGAGAAAGAAATATAGTATGAGCCCGAGCACCGCA 361
QY 388 TGTGCTGCTCCGCTGCGCGCCAGGACCTATGCTCAGCTAAATGATAGCCGATCCGGG 447
DB 362 TGTGCTGCTCCGCTGCGCGCCAGGACCTATGCTCAGCTAAATGATAGCCGATCCGGG 421
QY 448 ACACAGTTTGTGCCACATGTGCGGAGAAATCTACACAGGAGCTGGAATCTACCTGACCA 507
DB 422 ACACAGTTTGTGCCACATGTGCGGAGAAATCTACACAGGAGCTGGAATCTACCTGACCA 481
QY 508 TGTGCGAGCTGTGCGCGCCCTGTGACCCAGTGTGATGGGCTCGAGGAGATTGCCCGCTGCA 567
DB 482 TGTGCGAGCTGTGCGCGCCCTGTGACCCAGTGTGATGGGCTCGAGGAGATTGCCCGCTGCA 541
QY 568 CAAGCAAAACGAGACCGAGTCCGCTGCGGATGCGGCGGAAATGTTGTGTGCTGCTGGGCC 627
DB 542 CAAGCAAAACGAGACCGAGTCCGCTGCGGATGCGGCGGAAATGTTGTGTGCTGCTGGGCC 601
QY 628 TCGAGTGTACACATCGGAGCTACTTCTGACTGCCCGCTGGCACTGAAGCGGAGCTCA 587
DB 602 TCGAGTGTACACATCGGAGCTACTTCTGACTGCCCGCTGGCACTGAAGCGGAGCTTC 661
QY 688 AAGATGAAGTGGGAGGGTAAACACCTCGCTCCCTCGCAAGGCGAGGCGACTTCCAGA 747
DB 662 AAGATGAAGTGGGAGGGTAAACACCTCGCTCCCTCGCAAGGCGGCGACTTCCAGA 721

QY 748 ATACCTCTCCCGCCAGCGCCGCTGCCAGCCCGCCACACACAGGTGTGAG-AACCAAGGTCTG 806
DB 722 ATACCTCTCCCGCCAGCGCCGCTGCCAGCCCGCCACACACAGGTGTGAGAAACAGGCGCTG 781
QY 807 GTGGAGGAGCTCCAGGCACT-GCCAGTCCGACACAACT-CCAAAATCCATTAGAC 864
DB 782 GTGGAGGAGCTCCAGGCACTGGCCAGTCCGACACAACTGGCAAAATCCCTTAGAG 841
QY 865 CACTGCCCCCAGAGA 879
DB 842 CCTGGCCCCCAGAGA 856
AL564472 978 bp mRNA linear EST 16-FEB-2001
AL564472 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DM007YN10 3
LOCUS prime, mRNA sequence.
DEFINITION AL564472.1 GI:12914912
ACCESSION AL564472
VERSION AL564472
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 978)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..978
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DM007YN10"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com".
BASE COUNT 160 a 280 c 305 g 227 t 6 others
ORIGIN
Query Match 40.2%; Score 797.4; DB 9; Length 978;
Best Local Similarity 95.8%; Pred. No. 3.3e-170;
Matches 932; Conservative 5; Mismatches 21; Indels 15; Gaps 11;
QY 1011 GTTTCCTCCAGTATCCACTGGGCTCCCGCCAGCCCGGCTTTGGAGCAGGGGTGCCGCA 1070
DB 974 GTTTCCTCCAGTATCCACT-GGCTCCCGCCAGCCCGCAKTTT--GGAGGCGAGGTGCCGCA 918
QY 1071 CAGCAGAGTCTCTGGAGCTGACAGGAGCGCAGTTGGAACCGGGGAGCAGAGCCAG 1130
DB 917 CAGCAGAGTCTCT-GACCTGACCA-GGAGCGCAGTKGACCC--GGAGGAGAGACCA- 863
QY 1131 GTGGCCACCGGTACCAATGCAATTCATGTCACCGCGGGTCTATGACTATCACTGCGCAAC 1190
DB 862 GTGGCCACCGGTACCAATGCAATTCATGTCACCGCGGG--TCTATATCACT-GCAAC 806
QY 1191 ATCTACATCTACAAATGACAGTACTGGGGGACACCGGGTCTCTGGAGACCTCCAGCT 1250
|||||

Db 805 ATCTACATCTACAATGGACCAAGWACT - GGGGGACACACGGGTCTCTGGAGACCTCCCGAGCT 747

Qy 1251 ACCCCCGAACCTCCATACCCCATCCCGAAGAGGGGACCTGGCCCTCCCGGGCTCTCT 1310

Db 746 ACCCCCGAACCTCCATACCCCATCCCGAAGAGGG - ACCTGSCCTCCCGGGCTCTCT 689

Qy 1311 ACACCCCAACAGGAGAGTGGCAAGCTTTGGCACTTAGCGGAGACAGACACTGTGGTGC 1370

Db 688 ACACCCCAACAGGAGAGTGGCAAGCTTTGGCACTTAGCGGAGACAGACACTGTGGTGC 629

Qy 1371 ACACCCCTACAGGGGCCCCAAGAACCAATTTATCACCCTAGTACTGACTGAGTCTAG - 1429

Db 628 ACACCCCTACAGGGGCCCCAAGAACCAATTTATCACCCTAGTACTGACTGAGTCTAGT 569

Qy 1430 AAAAGGCAGAAAGAGGGGGGACAAAGGCACCTTCTCCCTTGAGGCTGCCCTGCCACGT 1489

Db 568 AAAAGGCAGAAAGAGGGGGGACAAAGGCACCTTCTCCCTTGAGGCTGCCCTGCCACGT 509

Qy 1490 GGGATTACAGGGGCTGAGTAGGGCCCGGGGAGAGAGCCCTAAAGGGATTAAAGGCTCA 1549

Db 508 GGGATTACAGGGGCTGAGTAGGGCCCGGGGAGAGAGCCCTAAAGGGATTAAAGGCTCA 449

Qy 1550 GACACCTCTGAGACAGTGGGCACTGGCTGAGTACGCTGCTCCACAGCACTCCCT 1609

Db 448 GACACCTCTGAGACAGTGGGCACTGGCTGAGTACGCTGCTCCACAGCACTCCCT 389

Qy 1610 ACTGCTGAGCAAACTGAGGCTCCCGGCAGACCCACCCCTGGGCTGCTCAGC 1669

Db 388 ACTGCTGAGCAAACTGAGGCTCCCGGCAGACCCACCCCTGGGCTGCTCAGC 329

Qy 1670 CTCAGGCAGACAGGGGACATGATACCACTGCTGCCCCACTACAGACCGCGCACCGGA 1729

Db 328 CTCAGGCAGACAGGGGACATGATACCACTGCTGCCCCACTACAGGCACCGCGCACCGGA 269

Qy 1730 GCAGGCAACCGAGGAGCGCCACACAGCTACCTGCAAGGAGCTACGSGGCCCTCTAAA 1789

Db 268 GCAGGCAACCGAGGAGCGCCACACAGCTACCTGCAAGGAGCTACGSGGCCCTCTAAA 209

Qy 1790 GGATTCGTGTGCTCATCCCCCAAGCTTCAGAGACCTTTGGGGTTCCACACTTCACGTGG 1849

Db 208 GGATTCGTGTGCTCATCCCCCAAGCTTCAGAGACCTTTGGGGTTCCACACTTCACGTGG 149

Qy 1850 ACTGAGTAGACCTGCGATGAGATGAATATAGGAGGAGCGTCTCTCCCTCCCTCC 1909

Db 148 ACTGAGTAGACCTGCGATGAGATGAATATAGGAGGAGCGTCTCTCCCTCCCTCC 89

Qy 1910 TAGAGGAGGAAAGGAGTCAATTAACAACCTAGGGGTTGGTAGGATTCCTAGGTATGG 1969

Db 88 TAGAGGAGGAAAGGAGTCAATTAACAACCTAGGGGTTGGTAGGATTCCTAGGTATGG 29

Qy 1970 GGAAGAGTTTGG 1982

Db 28 GGAAGAGTTTGG 16

RESULT 7

AL513838 Locus

DEFINITION AL513838 LTI_NFL006_P12 Homo sapiens cDNA clone CL0BA002ZF12 5

ACCESSION AL513838

VERSION AL513838.1 GI:12777332

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers
1. 993
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CL0BA002ZF12"
/clone_lib="LTI_NFL006_P12"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site.1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 184 a 348 c 278 g 180 t 3 others

ORIGIN

Query Match 40.1%; Score 794.8; DB 9; Length 993;
Best Local Similarity 97.9%; Pred. No. 1.3e-169;
Matches 805; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 78 CGGCGCTGGGGTGACATCGGCGCTGAGTCCCGTCCAGGCTCTGGGCTCGGGCAGCGCG 137

Db 5 CGGTCCGGAATTCCTCGGGATGCTGCTGAGTCCCGTCCAGGCTCTGGGCTCGGGCAGCGCG 64

Qy 138 CGCACCGCTCGCCAGGACGTGCGGGCTCTGCTCTTCTCCAGGCCCCACAGTTGCTGG 197

Db 65 CGCACCGCTCGCCAGGACGTGCGGGCTCTGCTCTTCTCCAGGCCCCACAGTTGCTGG 124

Qy 198 CCGCTGCGCGAGTGGCGCATGCTGCTGCTTGGGCGACCTCTGCCCCCGCGCTGCC 257

Db 125 CCGCTGCGCGAGTGGCGCATGCTGCTGCTTGGGCGACCTCTGCCCCCGCGCTGCC 184

Qy 258 TGGGGGCTCTGCTGCTGGGCTCTTGGGCTCTGCGAGCATGCGAGCGCCAGCGCGTG 317

Db 185 TGGGGGCTCTGCTGCTGGGCTCTTGGGCTCTGCGAGCATGCGAGCGCCAGCGCGTG 244

Qy 318 CCTCATATGCTCGGAGAACAGACCTGCGAGGACAGGAAAGAAATATATGAGGCC 377

Db 245 CCTCATATGCTCGGAGAACAGACCTGCGAGGACAGGAAAGAAATATATGAGGCC 304

Qy 378 CAGCACCGCATCTGCTGCTCCGCTGCCCGCGAGCACCTATGCTCAGCTAAATGTAGC 437

Db 305 CAGCACCGCATCTGCTGCTCCGCTGCCCGCGAGCACCTATGCTCAGCTAAATGTAGC 364

Qy 438 CGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAATTCCTACAACGAGCACTGGAAC 497

Db 365 CGCATCCGGGACACAGTTTGTGCCACATGTGCCGAGAATTCCTACAACGAGCACTGGAAC 424

Qy 498 TACCTGACCATCTGCCAGCTGTCGCGCCCTGTGACCCAGTGTATGGGCTCGAGGAGATT 557

Db 425 TACCTGACCATCTGCCAGCTGTCGCGCCCTGTGACCCAGTGTATGGGCTCGAGGAGATT 484

Qy 558 GCCCCTGACACAGCAACGAGGAGGAGTGGGCTGCCAGCGGGGATGTTCTGTGCT 617

Db 485 GCCCCTGACACAGCAACGAGGAGGAGTGGGCTGCCAGCGGGGATGTTCTGTGCT 544

Qy 618 GCCTGGGCGCTCGAGTGTACACACTGCGAGCTACTTTCTGACTGCCCGCTGGCACTGAA 677

Db 545 GCCTGGGCGCTCGAGTGTACACACTGCGAGCTACTTTCTGACTGCCCGCTGGCACTGAA 604

Qy 678 GCCGAGCTCAAAGATGAAGTTGGGAAGGTTAAACCACTCGCTGCCCTGCAAGGAGG 737

Db 605 GCCGAGCTCAAAGATGAAGTTGGGAAGGTTAAACCACTCGCTGCCCTGCAAGGAGG 664

Qy 738 CACTTCCAGATACCTCTCTCCCGCGCGCTGCCAGCGCCACACAGGTGTGAGAAC 797

Db 665 CACTTCCAGATACCTCTCTCCCGCGCGCTGCCAGCGCCACACAGGTGTGAGAAC 724

QY 798 CAAGTCTGCTGGAGGAGCTCCAGGCACTGCCAGTCCGACACACCAACCTGCAAAATCCA 857
 Db 725 CAAGTCTGCTGGAGGAGCTCCAGGCACTGCCAGTCCGACACCAACCTGCAAAATCCA 784

QY 858 TTAGAGCCACTGCCCCAGAGATGTGAGAGTGCCTGCTCAAG 899
 Db 785 TTAGAGCCACTGCCCCAGAGATGTGAGAGTGCCTGCTCAAG 826

RESULT 8
 B1908262
 LOCUS 603066987F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5216085 5',
 DEFINITION mRNA sequence.
 ACCESSION B1908262
 VERSION B1908262
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 816)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL).
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11542 row: n column: 22
 High quality sequence stop: 813.
 Location/Qualifiers
 1..816
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5216085"
 /clone_lib="NIH_MGC_118"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dr primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."

BASE COUNT 150 a 299 c 231 g 136 t
 ORIGIN

Query Match 40.1%; Score 794.2; DB 13; Length 816;
 Best Local Similarity 99.5%; Pred No. 1.6e-169;
 Matches 807; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 63 CGGCTGGCCGCTCCGCGCCCTGGGTGCACATCGGCCCTGATCGCCGTCGCCAGGCTCTG 122
 Db 1 CGGCTGGCCGCTCCGCGCCCTGGGTGCACATCGGCCCTGATCGCCGTCGCCAGGCTCTG 60

QY 123 GGTCTGGGAGCGCGCGCCGCTCCGAGGAGCTGCGGCTCTCTGCTTCTCTCCAGG 182
 Db 61 GGTCTGGGAGCGCGCGCCGCTCCGAGGAGCTGCGGCTCTCTGCTTCTCTCCAGG 120

QY 183 CCCCCACGTTGCTGGCGCGCTGCGCGAGTGGCGGCATCTCTGCTGGGCGCACCTCT 242
 Db 121 CCCCCACGTTGCTGGCGCGCTGCGCGAGTGGCGGCATCTCTGCTGGGCGCACCTCT 180

QY 243 GCCCGCGCTGGCGCTCTGCTGGGCTCTTCTGGGCTCTTGGCAGCATCG 302
 Db 181 GCCCGCGCTGGCGCTCTGCTGGGCTCTTCTGGGCTCTTGGCAGCATCG 240

QY 303 CAGCCCCAGCGGTGCTCCATATGCTGCGAGAACCAACCTGAGGAGCCAGGAAAG 362
 Db 241 CAGCCCCAGCGGTGCTCCATATGCTGCGAGAACCAACCTGAGGAGCCAGGAAAG 300

QY 363 GAATACTATGAGCCCCAGCACCGCATCTGCTGCTCCGCTGCCGCCAGCCTATGTC 422
 Db 301 GAATACTATGAGCCCCAGCACCGCATCTGCTGCTCCGCTGCCGCCAGCCTATGTC 360

QY 423 TCAGCTAAATGTAGCGCATTCGCGGACACAGTTTGTGCCACATGTGCCGAGAAATTCCTAC 482
 Db 361 TCAGCTAAATGTAGCGCATTCGCGGACACAGTTTGTGCCACATGTGCCGAGAAATTCCTAC 420

QY 483 AACGAGCACTGGAACCTGACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTGATG 542
 Db 421 AACGAGCACTGGAACCTGACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTGATG 480

QY 543 GGCCTCGAGGAGATTCGCCCTCGACAAGCAACGAAGACCCAGTGGCGCTGCCAGCCG 602
 Db 481 GGCCTCGAGGAGATTCGCCCTCGACAAGCAACGAAGACCCAGTGGCGCTGCCAGCCG 540

QY 603 GGAATGTTCTGTGCTGCTGCGCTGCGCCCTCGAGTGTACACACTGCGAGCTACTTTCTGACTGC 662
 Db 541 GGAATGTTCTGTGCTGCTGCGCTGCGCCCTCGAGTGTACACACTGCGAGCTACTTTCTGACTGC 600

QY 663 CGCCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTTAACCAACCTGCGTC 722
 Db 601 CGCCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGGTTAACCAACCTGCGTC 660

QY 723 CCTCTCAAGCGCAGGCACTTCCAGAAATCTCTCTCCCGCAGCGCCGCTGCCAGCCCCAC 782
 Db 661 CCTCTCAAGCGCAGGCACTTCCAGAAATCTCTCTCCCGCAGCGCCGCTGCCAGCCCCAC 719

QY 783 ACCAGTGTGGAACAACAAGTCTGTGGAGGAGCTTCCAGGCACTGCCAGTCCGACACACA 842
 Db 720 ACCAGTGTGGAACAACAAGTCTGTGGAGGAGCTTCCAGGCACTGCCAGTCCGACACACA 779

QY 843 ACCTCAAAAATCCATTAGAGCCACTGCCCC 873
 Db 780 AACTCAAAAATCCATTAGAGCCACTGCCCC 810

RESULT 9
 B1768812
 LOCUS 603057713F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5207139 5',
 DEFINITION mRNA sequence.
 ACCESSION B1768812
 VERSION B1768812.1 GI:15760390
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 867)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11519 row: j column: 04
 High quality sequence stop: 818.
 Location/Qualifiers

FEATURES

```

source
1. .867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5207139"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/Note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 171 a 306 c 248 g 141 t 1 others
ORIGIN

Query Match 39.6%; Score 785.6; DB 13; Length 867;
Best Local Similarity 98.1%; Pred. No. 1.5e-167;
Matches 794; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 91 CACATCGCCCTGAGTCCCGTCCAGGCTCTGGGCTCGGCAGCCGCCACCGCTGCC 150
DB 1 CACATCGCCCTGAGTCCCGTCCAGGCTCTGGGCTCGGCAGCCGCCACCGCTGCC 60
QY 151 CAGGACGTGGGCTCTGCTTCTCCAGGCCCCACGTTGTCGCCGCTGGCCGAG 210
DB 61 CAGGACGTGGGCTCTGCTTCTCCAGGCCCCACGTTGTCGCCGCTGGCCGAG 120
QY 211 TGGCGCCATGCTCTGCTTGGGCACTCTGCCCCGCGCTGGCTGGGGCTCTGG 270
DB 121 TGGCGCCATGCTCTGCTTGGGCACTCTGCCCCGCGCTGGCTGGGGCTCTGG 180
QY 271 TGCTGGGCTCTTCCGGCTCTGTCAGCATCGACCCCGAGGGGTGCTCCATATCGT 330
DB 181 TGCTGGGCTCTTCCGGCTCTGTCAGCATCGACCCCGAGGGGTGCTCCATATCGT 240
QY 331 CGGAGAACCATGTCAGGACCCAGGAAGAAAGTAATCTATGACCCCGAGCCGATCT 390
DB 241 CGGAGAACCATGTCAGGACCCAGGAAGAAAGTAATCTATGACCCCGAGCCGATCT 300
QY 391 GCTGCTCCGCTGCGCCGAGCACTATGCTCAGCTAAATGATAGCCGATCGGGACA 450
DB 301 GCTGCTCCGCTGCGCCGAGCACTATGCTCAGCTAAATGATAGCCGATCGGGACA 360
QY 451 CAGTTTGTGCACATGTGCCGAGAAATTCCTACACGAGCACTGGAATCTGACCATCT 510
DB 361 CAGTTTGTGCACATGTGCCGAGAAATTCCTACACGAGCACTGGAATCTGACCATCT 420
QY 511 GCCAGCTGCGCCGCTGTGACCCAGTGTAGGGCTCGAGGAGATGCCCCCTGCACAA 570
DB 421 GCCAGCTGCGCCGCTGTGACCCAGTGTAGGGCTCGAGGAGATGCCCCCTGCACAA 480
QY 571 GCAACGGAAGACCCAGTGGCGCTGCCAGCGGAAATGTTCTGTGCTGGGCGCTCG 630
DB 481 GCAACGGAAGACCCAGTGGCGCTGCCAGCGGAAATGTTCTGTGCTGGGCGCTCG 540
QY 631 AGTGTACACATGCGAGTACTTTCTGACTGCCCGCTGGGCACTGAAGCCGAGCTCAAG 690
DB 541 AGTGTACACATGCGAGTACTTTCTGACTGCCCGCTGGGCACTGAAGCCGAGCTCAAG 600
QY 691 ATGAAGTTGGGAAGGTAACAAACCATGCGTCCGCTCGAAGGCAAGGCACTTCCAGAATA 750
DB 601 ATGAAGTTGGGAAGGTAACAAACCATGCGTCCGCTCGAAGGCAAGGCACTTCCAGAATA 660
QY 751 CCTCTCCCCAGGCGCGCTGCCAGCCCCACACAGAGTGTGAGAACCAAGGTCTGTGG 810
DB 661 CCTCTCCCCAGGCGCGCTGCCAGCCCCACACAGAGTGTGAGAACCAAGGTCTGTGG 720
QY 811 AGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAATCCATTTAGAGCCACTGC 870
DB 811 AGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAATCCATTTAGAGCCACTGC 870

```

Db 721 AGGAGCTCCAGGCACTGCCAGTCCGACACGAACTGCAAAATCCATTAGAGCCAACTG 780
 QY 871 CCCAGAGATGTCAGGATCGCTGCTCAAG 899
 Db 781 CCCAGAGATGTCAGAACTGGCCTCAAG 809

RESULT 10
 BI906190 840 bp mRNA linear EST 16-OCT-2001
 LOCUS 603062686F1.NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5211812 5',
 DEFINITION mRNA sequence.
 ACCESSION BI906190
 VERSION BI906190.1 GI:16168845
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 840)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11531 row: 1 column: 21
 High quality sequence stop: 776.
 Location/Qualifiers
 1. .840
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5211812"
 /clone_lib="NIH_MGC_118"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."
 BASE COUNT 155 a 300 c 246 g 138 t 1 others

Query Match 38.1%; Score 755.4; DB 13; Length 840;
 Best Local Similarity 97.4%; Pred. No. 1e-160;
 Matches 821; Conservative 0; Mismatches 16; Indels 6; Gaps 5;

QY 47 GGGAGCCTTGAGGCGCGGCTGGCGCTCCCGGCTGGGGTGCACATCGGCCCTGAGT 106
 DB 1 GGGAGCCTTGAGGCGCGGCTGG-CGCTCCGCGCTGGGGTGCACATCGGCCCTGAGT 59
 QY 107 CCGGTCCAGGCTCTGGGCTCGGCGACCGCGCCACCGCTGCCAGGAGTCTGGGGCTC 166
 DB 60 CCGGTCCAGGCTCTGGGCTCGGCGACCGCGCCACCGCTGCCAGGAGTCTGGGGCTC 119
 QY 167 CTGCTTCTCCAGGCGCCACGTTGTCGGCCCTGGCGAGTGGCCGATGCTCT 226
 DB 120 CTGCTTCTCCAGGCGCCACGTTGTCGGCCCTGGCGAGTGGCCGATGCTCT 179
 QY 227 GCCTTGGGCGACCTCTGCCCCCGGCTGGCTGGGGGCTCTGTGTGGGCTCTTCGG 286
 DB 180 GCCTTGGGCGACCTCTGCCCCCGGCTGGCTGGGGGCTCTGTGTGGGCTCTTCGG 239

QY	287	GCTCTGGAGCATCGAGCCCGAGCGGTGCTTCCATATGCTCGGAGAACAGACCTG	346
Db	240	GCTCTGGAGCATCGAGCCCGAGCGGTGCTTCCATATGCTCGGAGAACAGACCTG	299
QY	347	CAGGACACAGGAAAGAAATATATATGAGCCCGAGCCCGAGCAGCTGCTCGCGTGC	406
Db	300	CAGGACACAGGAAAGAAATATATATGAGCCCGAGCCCGAGCAGCTGCTCGCGTGC	359
QY	407	GCAGGACCATATGCTCAGTAAATGTAGCCGATCCGAGCAGCTGCTCGCGTGC	466
Db	360	GCAGGACCATATGCTCAGTAAATGTAGCCGATCCGAGCAGCTGCTCGCGTGC	419
QY	467	TGCCGAGAAATTCCTACAAAGAGCAGCTGGAATACCTGACCATCTGCCAGCTGTGCCGCC	526
Db	420	TGCCGAGAAATTCCTACAAAGAGCAGCTGGAATACCTGACCATCTGCCAGCTGTGCCGCC	479
QY	527	CTGTGACCCAGTATGGGCTCGAGGAGATGCCCCCTGCACAAAGCAAGCAAGACCCA	586
Db	480	CTGTGACCCAGTATGGGCTCGAGGAGATGCCCCCTGCACAAAGCAAGCAAGACCCA	539
QY	587	GTCCCGCTGCGAGCGGAAATGTTCTGCTGCTGGCCCTCGAGTGTACACACTGCGA	646
Db	540	GTCCCGCTGCGAGCGGAAATGTTCTGCTGCTGGCCCTCGAGTGTACACACTGCGA	599
QY	647	GCTACTTTCTGACTGCGCGCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGG	706
Db	600	GCTACTTTCTGACTGCGCGCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAAGG	659
QY	707	TAAACACCACTGCTGCTGCAAGGAGGCACTTCCAGATACCTCTCCCGCCAGCGC	766
Db	660	TAAACACCACTGCTGCTGCAAGGAGGCACTTCCAGATACCTCTCCCGCCAGCGC	718
QY	767	CCGCTGCCAGCCCGAGGAGTGTGAGACCAAGGCTGCTGGAGGCACTCCAGGCAC	826
Db	719	CCGCTGCCAG-GCCACACCAAGGTTGAGAACCAAGGCTGCTGGAGGCACTCCAGGCAC	777
QY	827	TGCCAGTCCGACA-CAAGCTGCAAAATCCATTAGAGCCACTGCCCGCCAGAGATGTGAG	886
Db	778	TGCCAGTCCGACA-CAAGCTGCAAAATCCATTAGAGCCACTGCCCGCCAGAGATGTGAG	834
QY	887	ATC 889	
Db	835	AAC 837	
RESULT 11			
BI821009			
LOCUS	603035664F1	NIH_MGC_115	Homo sapiens cDNA clone IMAGE:5176935 5',
DEFINITION	mRNA sequence.		
ACCESSION	BI821009		
VERSION	BI821009.1	GI:15932559	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	NIH-MGC http://mgi.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgaabs-r@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM1440 row: 0 column: 16 High quality sequence stop: 758.		

FEATURES	source	Location/Qualifiers
		1. .954
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="IMAGE:5176935"
		/clone_lib="NIH_MGC_115"
		/lab_host="DH10B"
		/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."
BASE COUNT	212 a	313 c 281 g 148 t
ORIGIN		
	Query Match	37.7%; Score 746.6; DB 13; Length 954;
	Best Local Similarity	93.3%; Pred. No. 1e-158;
	Matches 861; Conservative	0; Mismatches 44; Indels 18; Gaps 7;
QY	885	GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGGAGCCCAATCTGTAGCTGGAGCTGG 944
Db	14	GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGGAGCCCAATCTGTAGCTGGAGCTGG 73
QY	945	GAGCTCCGAAAGGCCATCCATCTTCCCTGACTTGTGTACAGCCACTGTCTACCCATTCT 1004
Db	74	GAGCTCCGAAAGGCCATCCATCTTCCCTGACTTGTGTACAGCCACTGTCTACCCATTCT 133
QY	1005	GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGCCAGTTTGGAGGCGAGGGGTG 1064
Db	134	GGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCCCGCCAGTTTGGAGGCGAGGGGTG 193
QY	1065	CCGCAACAGCAGAGTCTCTTGGACCTGACGAGGAGCGCGAGTTTGAACCCGGGAGCAG 1124
Db	194	CCGCAACAGCAGAGTCTCTTGGACCTGACGAGGAGCGCGAGTTTGAACCCGGGAGCAG 253
QY	1125	AGCCAGGTGGCCACGGTACCAATGGCATTCATGTACCGCGGGGTCTATGACTATCACT 1184
Db	254	AGCCAGGTGGCCACGGTACCAATGGCATTCATGTACCGCGGGGTCTATGACTATCACT 313
QY	1185	GGCAACATCTACATCTACAATGGACCACTACTGGGGGAGACACCGGGTCTCTGGAGACCTC 1244
Db	314	GGCAACATCTACATCTACAATGGACCACTACTGGGGGAGACACCGGGTCTCTGGAGACCTC 373
QY	1245	CCAGCTACCCCGCAACCTCCATACCCCATTCGCCGAGAGGGGAGCCCTGGCCCTCCCGGG 1304
Db	374	CCAGCTACCCCGCAACCTCCATACCCCATTCGCCGAGAGGGGAGCCCTGGCCCTCCCGGG 433
QY	1305	CTCTTACACCCACAGGAGATGGCAAGCTTGGCACCTAGCGGAGACAGACGACTGT 1364
Db	434	CTCTTACACCCACAGGAGATGGCAAGCTTGGCACCTAGCGGAGACAGACGACTGT 493
QY	1365	GGTGCCACACCTCTAACAGGGGCCCAAGAACCAATTTATCACCCTGACTGACTGAGT 1424
Db	494	GGTGCCACACCTCTAACAGGGGCCCAAGAACCAATTTATCACCCTGACTGACGGAGT 553
QY	1425	CTGAGAAAGGCGAGAGAGGGGGGACAAAGGCGACCTTCTCCCTTGGAGCTGCCCTGCC 1484
Db	554	CTGAGAAAGGCGAGAGAGGGGGGAC - AGGGCGACCTTCTCCCTTGGAGCTGCCCTGCC 612
QY	1485	CAGTGGGATTACAGGGGCTGAGTAGGGCCCGGGGAGAGAGCCCTTAAGGATTAAAG 1544
Db	613	CAGTGGGATTACAGGGGCTGAGTAGGGCCCGGGGAGAGAGCCCTTAAGGATTAAAG 672
QY	1545	GCTCAGACACTCTTGAGAGAGGTGGCAGCTGGCTGGG - -TACGGTCCCTCCACAGGAC 1602
Db	673	GCTCAGACACTCTTGAGAGAGGTGTGCATGTGGCTGGGGTACGGTTGCCCTCCACAGGAC 732
QY	1603	TCTCCCTACTGCC -TGAGCAAACTGTAG -GCCTCCCGGAGACCCACCCCTCCCTGGGG 1660

```

Db 733 TCTCCCTACTGAGCAACCTGAGTCTCCCGGAGACCAACCCACCCCTGGG 792
QY 1661 CTGCTCAGCCTCAGSCAGGACGGGACATG-----ATACCAACTGCTGCCA 1709
Db 793 GGCCTGCTCAGCCTCCAGGACGAGGACGACATGATGATACCAAGCTGCTGCCA 852
QY 1710 CTACAGCAGCCGACGACGAGGACGACGACGAG-GGGAGCCGCCACAC-GGTCACCTGCAA 1767
Db 853 CTAAGGCAACGCGGACGAGCAGGACGAGCAGGATGGAGCGCCACACGCGGTCACTGCAA 912
QY 1768 GGAGCTACGCGGCGCCCTTAAG 1790
Db 913 GGAGCTACGCGGCGCCCTCCTAAGG 935

RESULT 12
BO957322 895 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8777754 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384424
5', mRNA sequence.
ACCESSION BO957322
VERSION BO957322.1 GI:22372800
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2579 row: g column: 17
High quality sequence stop: 645.
FEATURES
Location/Qualifiers
1..895
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6384424"
/clone_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC Library.
BASE COUNT 190 a 291 c 252 g 159 t 3 others
ORIGIN

Query Match 36.88; Score 729; DB 14; Length 895;
Best Local Similarity 98.7%; Pred. No. le-154;
Matches 754; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 885 GGATCGCTGCTCAAGAGCGCTCCGACGGAGAGGACCCCAATCTGTAGCTGGAAGCTGG 944
Db 132 GGATCGCTGCTCAAGAGCGCTCCGACGGAGAGGAGGACCCCAATCTGTAGCTGGAAGCTGG 191
QY 945 GAGCCTCCGAGGCGCCATCCATCTCCCTGACTTGTGTACAGGCACATGCTACCCATTCT 1004
Db 192 GAGCCTCCGAGGCGCCATCCATCTCCCTGACTTGTGTACAGGCACATGCTACCCATTCT 251

```

```

QY 1005 GGAGATGTTTCCCCAGTATCCACTGGGTCCCGCAGCCCCAGTTTGGAGCGAGGGTG 1064
Db 252 GGAGATGTTTCCCCAGTATCCACTGGGTCCCGCAGCCCCAGTTTGGAGCGAGGGTG 311
QY 1065 CCGCAACAGCAGAGTCTCTGGACCTGACAGGAGGCGCAGTTGGAAACCCGGGAGCAG 1124
Db 312 CCGCAACAGCAGAGTCTCTGGACCTGACAGGAGGCGCAGTTGGAAACCCGGGAGCAG 371
QY 1125 AGCCAGTGGCCCGCAGTACCAATGGGATTCATGTCACCGGCGGTCTATGACTATCACT 1184
Db 372 AGCCAGTGGCCCGCAGTACCAATGGGATTCATGTCACCGGCGGTCTATGACTATCACT 431
QY 1185 GGCAACATCTACATCTACAATGGACCACTACTGGGGGACCCACCGGCTCTTGGAGACCTC 1244
Db 432 GGCAACATCTACATCTACAATGGACCACTACTGGGGGACCCACCGGCTCTTGGAGACCTC 491
QY 1245 CCAGCTACCCCGCAACCTCCATACCCCATTCGCCAAGAGGGGACCTGGCCCTCCCGGG 1304
Db 492 CCAGCTACCCCGCAACCTCCATACCCCATTCGCCAAGAGGGGACCTGGCCCTCCCGGG 551
QY 1305 CTCCTACACCCCGCAGGAGATGGCAAGGCTTGGCACCTAGCGAGAGACAGACACTGT 1364
Db 552 CTCCTACACCCCGCAGGAGATGGCAAGGCTTGGCACCTAGCGAGAGACAGACACTGT 611
QY 1365 GGTGCCACACCTCTAACAGGGGCCCAAGGAACCAATTTATCACCCATGACTGACTGAGT 1424
Db 612 GGTGCCACACCTCTAACAGGGGCCCAAGGAACCAATTTATCACCCATGACTGACTGAGT 671
QY 1425 CTGAGAAAAGCAGAGAAGGGGGGACAGAGGACCTTCTCCCTTGGAGGTGCCCTGCC 1484
Db 672 CTGAGAAAAGCAGAGAAGGGGGGACAGAGGACCTTCTCCCTTGGAGGTGCCCTGCC 731
QY 1485 CACGTGGATTTCACAGGGGCTGAGTAGGGCCGCGGAAGCAGAGCCCTAAGGATTAAAG 1544
Db 732 CACGTGGATTTCACAGGGGCTGAGTAGGGCCGCGGAAGCAGAGCCCTAAGGATTAAAG 790
QY 1545 GCTCAGACACCTCTGAGAGCAGGTGGGCATCTGGTGGGTACGGTGCCTCCACAGGACT 1603
Db 791 GCTCAGACACCTCTGAGAGCAGGTGGGCATCTGGTGGGTACGGTGCCTCCACAGGACT 850
QY 1604 CTCCTACTGCTGAGCAAAACCTGAGGCTCCCGGACAGACCCAC 1647
Db 851 CTCCTACTGCTGAGCAAAACCTGAGGCTCCCGGACAGACCCAC 894

```

RESULT 13

BI911211

LOCUS

DEFINITION

BI911211.1 GI:16174826

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI911211 891 bp mRNA linear EST 16-OCT-2001
603062959F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:521047 5',
mRNA sequence.

BI911211
BI911211.1 GI:16174826

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 891)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM1532 row: f column: 16

High quality sequence start: 26


```

Db 181 CTGCTGCTGGCCCTCTTCTGGGCTCTCTGCGACATCGAGCCCGAGGGTGCCTCCATAT 240
Qy 327 GCGTCGGAGAACACAGACTCGCAGGAGACAGAAAAGAAATACTATGAGCCCGACACCGC 386
Db 241 GCGTCGGAGAACACAGACTCGCAGGAGACAGAAAAGAAATACTATGAGCCCGACACCGC 300
Qy 387 ATCTGCTGCTCCGCTGCGCCCGCCAGGACCTATCTCTCAGCTAAATGTAGCCGATCCGG 446
Db 301 ATCTGCTGCTCCGCTGCGCCCGCCAGGACCTATCTCTCAGCTAAATGTAGCCGATCCGG 360
Qy 447 GACACAGTTGTGTCACATGTCGCGAGAAATCTTCTACACGAGCACTGGAATCTGACCT 506
Db 361 GACACAGTTGTGTCACATGTCGCGAGAAATCTTCTACACGAGCACTGGAATCTGACCT 420
Qy 507 ATCTGCCAGCTGTGTCGCGCCCTGTGACCCAGTGTGCGGCTCGAGGAGATTGCCCCCTGC 566
Db 421 ATCTGCCAGCTGTGTCGCGCCCTGTGACCCAGTGTGCGGCTCGAGGAGATTGCCCCCTGC 480
Qy 567 ACAAGCAAAAGGAGACAGCTGCGCTGCGAGCCGCGGAAATGTTCTGTGCTGCGTGGGC 626
Db 481 ACAAGCAAAAGGAGACAGCTGCGCTGCGAGCCGCGGAAATGTTCTGTGCTGCGTGGGC 540
Qy 627 CTCGAGTGTACACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAGCCGAGCTC 686
Db 541 CTCGAGTGTACACACTGCGAGCTACTTCTGACTGCGCGCTGCGACTGAGCCGAGCTC 600
Qy 687 AAAGATGAAGTTGGGAAGGTTA-ACAACCACTGCGTCCCTGCAAGGCA-GGGCACTTCC 744
Db 601 AAAGATGAAGTTGGGAAGGTTAGCAACCACTGCGTCCCTGCAAGGCAAGGCGGCACTTCC 660
Qy 745 AGAATACCTCTCCCGCCAGCGCCGCTGCGAGCCCGCCAGCCAGAGTGTGAGAACCAAGTC 804
Db 661 AGAATACCTCTCTCCCGCCAGCGCCGCTGCGAGCCCGCCAGCCAGAGTGTGAGAACCAAGTC 720
Qy 805 TGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAATCCATTAGAGC 864
Db 721 TGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAATCCATTAGAGC 780
Qy 865 CACTGCC 872
Db 781 CATGCC 788

RESULT 15
Bi761249
LOCUS 603043491f1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183638 5',
DEFINITION mRNA sequence.
ACCESSION Bi761249
VERSION Bi761249.1 GI:15752827
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 881)
NIH-MGC http://imgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arranged by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11458 row: f column: 23
High quality sequence stop: 815.
Location/Qualifiers
1. .881

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5183638"
/lab_lib="NIH_MGC_116"
/lab_host="DH10B"
/notes="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SpOrf6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."
BASE COUNT 160 a 315 c 255 g 151 t
ORIGIN

```

```

Query Match 36.0%; Score 714.4; DB 13; Length 881;
Best Local Similarity 97.6%; Pred. No. 2e-151;
Matches 810; Conservative 0; Mismatches 11; Indels 9; Gaps 8;

Qy 60 GCCGGGCTGGCGCTCCCGGCTGGGTGTCACATCGGCGCTGAGTCCCGTCCAGGCT 119
Db 8 GGCCTGGCTGGCGCTCCCGGCTGGGTGTCACATCGGCGCTGAGTCCCGTCCAGGCT 67
Qy 120 CTGGGCTCGGGCAGCC-GCCGCCACCGCTGCCAG-GACGTGCGGGCTCTCTGCTTCTC 177
Db 68 CTGGGCTCGGGCAGCTGCGGCCACCGCTGCCAGTGAAGTGGGCTCTCTGCTTCTC 127
Qy 178 CCAGGCCCCACAGTGTGTGCGCGCTGGCGAGTGGCGGCATGCTCTGCTGCTG-GGCC 236
Db 128 CCAGGCCCCACAGTGTGTGCGCGCTGGCGAGTGGCGGCATGCTCTGCTTGTGGCC 187
Qy 237 ACCTCTGCCCGCGCTGGCTGGGGGCTCTGTGTGCTGGGCTCTTTCGGGCTCTGSCA 296
Db 188 ACCTCTGCCCGCGCTGGCTGGGGGCTCTGTGTGCTGGGCTCTTTCGGGCTCTGSCA 247
Qy 297 GCATCGAGCCCCAGGCGGTGCTTCCATATGCGTGGAGAACCAAGACTCGAGGACAG 356
Db 248 GCATCGAGCCCCAGGCGGTGCTTCCATATGCGTGGAGAACCAAGACTCGAGGACAG 307
Qy 357 GAAAGAAATATATGAGCCCCAGCACCAGCATCTGCTGCCGCTGCCCGCCAGGACAC 416
Db 308 GAAAGAAATATATGAGCCCCAGCACCAGCATCTGCTGCCGCTGCCCGCCAGGACAC 367
Qy 417 TATGCTCAGCTAAATGTCAGCGCATCGGAGACAGTTTGTGCCACATGTGCCGAGAAT 476
Db 368 TATGCTCAGCTAAATGTCAGCGCATCGGAGACAGTTTGTGCCACATGTGCCGAGAAT 427
Qy 477 TCCTACAACGAGCACTGGAATACCTGACCATCTGCCAGCTGTGCCGCTGCCCGCTGTGACCCA 536
Db 428 TCCTACAACGAGCACTGGAATACCTGACCATCTGCCAGCTGTGCCGCTGCCCGCTGTGACCCA 487
Qy 537 GTGATGGGCTCGAGGAGATTGCCCCCTGCACAGCAAAAGGAGAGAGCCAGTGGCGTGC 596
Db 488 GTGATGGGCTCGAGGAGATTGCCCCCTGCACAGCAAAAGGAGAGAGCCAGTGGCGTGC 547
Qy 597 CAGCGGGAATGTTCTGTGCTGCTGGGCTCGAGTGTACACACTCGAGCTACTTCT 656
Db 548 CAGCGGGAATGTTCTGTGCTGCTGGGCTCGAGTGTACACACTCGAGCTACTTCT 607
Qy 657 GACTGCCCGCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGA-GGGTAACACCA 715
Db 608 GACTGCCCGCTGGCACTGAAGCCGAGCTCAAGATGAAGTTGGGAACGGGTAAACACCA 667
Qy 716 CTGGTCCCTGCAAGCAGGCACTTCCAGATACCTCTCCCGCCAGCCCGCTGCGCA 775
Db 668 CTGGTCCCTGCAAGCAGGCACTTCCAGATACCTCTCCCGCCAGCCCGCTGCGCA 727
Qy 776 GCCCCACACAGGTGTGAGAAC-AAGGTCTGGTGGAG-GCAGCTCCAGGC-ACTGCCCA 832
Db 776 GCCCCACACAGGTGTGAGAAC-AAGGTCTGGTGGAG-GCAGCTCCAGGC-ACTGCCCA 832

```

Db 728 GCGCCGCAACAGGTGTGAGAACCAAAAGTCTGGTGAGCGCAGCTCCAGGCAACTGCCCA 787
QY 833 GTCCGACACAACTGCAAAA--TCCATTAGAGCCCACTGCCCCACAGAGAT 880
|||||
Db 788 GTCCGACACAACTGCAAAAAGTCCATTAGGAGCCCACTGCCCCACAGAGAT 837
|||||

Search completed: April 15, 2003, 16:26:18
Job time : 2303.32 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:11:20 ; Search time 63.8651 Seconds
(without alignments)
9517.464 Million cell updates/sec

Title: US-09-917-372-2

Perfect score: 1982

Sequence: 1 gccccgcggccagctcgctc.....ggtatgggggaagatttgg 1982

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2.6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2.6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2.6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2.6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2.6/ptodata/1/ina/PTUS_COMB.seq:*
- 6: /cgn2.6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	3.6	1557	1 US-08-385-229-3	Sequence 3, Appli
2	72	3.6	1641	1 US-08-385-229-1	Sequence 1, Appli
3	72	3.6	1641	2 US-08-650-000-1	Sequence 1, Appli
4	72	3.6	1641	6 5395760-1	Patent No. 5395760
5	72	3.6	2224	4 US-08-477-347-2	Sequence 2, Appli
6	72	3.6	2224	4 US-08-476-862-1	Sequence 1, Appli
7	72	3.6	3683	4 US-09-844-634-3	Sequence 3, Appli
8	62.6	3.2	691	1 US-08-266-080B-12	Sequence 12, Appl
9	62.6	3.2	691	5 PCT-US95-05423-12	Sequence 12, Appl
10	62.6	3.2	705	4 US-09-326-394-3	Sequence 3, Appli
11	62.6	3.2	705	4 US-09-580-235-1	Sequence 1, Appli
12	62.6	3.2	705	4 US-09-580-235-3	Sequence 3, Appli
13	62.6	3.2	705	4 US-09-580-235-5	Sequence 5, Appli
14	62.6	3.2	705	4 US-09-580-235-7	Sequence 7, Appli
15	62.6	3.2	705	4 US-09-580-181-1	Sequence 1, Appli
16	62.6	3.2	705	4 US-09-580-181-3	Sequence 3, Appli
17	62.6	3.2	705	4 US-09-580-181-5	Sequence 5, Appli
18	62.6	3.2	705	4 US-09-580-181-7	Sequence 7, Appli
19	62.6	3.2	705	4 US-09-102-530-1	Sequence 1, Appli
20	62.6	3.2	705	4 US-09-102-530-3	Sequence 3, Appli
21	62.6	3.2	705	4 US-09-102-530-5	Sequence 5, Appli
22	62.6	3.2	705	4 US-09-102-530-7	Sequence 7, Appli
23	62.4	3.1	579	4 US-09-146-950-3	Sequence 3, Appli
24	62.4	3.1	591	4 US-09-146-950-19	Sequence 19, Appl
25	62.4	3.1	1596	4 US-09-146-950-17	Sequence 17, Appl
26	62.4	3.1	1724	4 US-08-509-024-1	Sequence 1, Appli
27	62.4	3.1	1724	4 US-09-333-279-1	Sequence 1, Appli

28	62.4	3.1	1724	5 PCT-US96-12374-1	Sequence 1, Appli
29	62.4	3.1	1929	4 US-09-146-950-1	Sequence 1, Appli
30	62.4	3.1	4622	4 US-08-509-024-6	Sequence 6, Appli
31	62.4	3.1	4622	4 US-09-333-279-6	Sequence 6, Appli
32	56	2.8	1391	3 US-08-996-139-3	Sequence 3, Appli
33	56	2.8	1391	4 US-08-995-659-3	Sequence 3, Appli
34	56	2.8	1391	4 US-09-215-649A-3	Sequence 3, Appli
35	56	2.8	1391	4 US-09-577-780-3	Sequence 3, Appli
36	56	2.8	3115	3 US-08-996-139-1	Sequence 1, Appli
37	56	2.8	3115	4 US-08-995-659-1	Sequence 1, Appli
38	56	2.8	3115	4 US-09-215-649A-1	Sequence 1, Appli
39	56	2.8	3115	4 US-09-577-780-1	Sequence 1, Appli
40	56	2.8	3136	3 US-08-996-139-5	Sequence 5, Appli
41	56	2.8	3136	4 US-09-435-296-3	Sequence 3, Appli
42	56	2.8	3136	4 US-08-995-659-5	Sequence 5, Appli
43	56	2.8	3136	4 US-09-215-649A-5	Sequence 5, Appli
44	56	2.8	3136	4 US-09-577-780-5	Sequence 5, Appli
45	52.6	2.7	519	2 US-08-249-189-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-385-229-3
; Sequence 3, Application US/08385229
; Patent No. 5605690
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Cindy A.
; APPLICANT: Smith, Craig A.
; TITLE OF INVENTION: Method of Treating TNF-Dependent
; TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,229
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,236
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wight, Christopher L.
; REGISTRATION NUMBER: 31,680
; REFERENCE/DOCKET NUMBER: 2503
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 587-0606
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1557 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: INFR/FC Fusion Protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1557


```

; FILING DATE: 09-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/547,750
; FILING DATE: 02-JULY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/505,441
; FILING DATE: 06-APRIL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: SYNE200C5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 691 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-266-080B-12

Query Match
Best Local Similarity 3.2%; Score 62.6; DB 1; Length 691;
Matches 165; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

QY 300 TGCAGCCCGCCAGCGGTGCTCCATATGCTCGGAGAACACAGACCTGCGAGGACACAGGAA 359
Db 139 TTGCCCGCCCGCCAGGTGCGCATTTACACCTACGCTCCGGAACCGGGTCTACCTGCGGCTC 198
QY 360 AAGGAATATATGAGCCCGCCAGCAGCATCTGCTGCCGTGCGCGCCGCGCAGGACCTAT 419
Db 199 AGAGAATACTATGACACAGACAGCTCAGATGTGCTGAGCAAGTGTGCTGCGCGGCGCAACAT 258
QY 420 GTCTCAGCTAAATGTAGCGCATCCGGGACACAGTTTGGCCACATGTGCGCGAGATTC 479
Db 259 GCAAAAGTCTTCTGTACCAAGACCTCGGACACCGGTGTGTGACTCTCTGTGAGGACAGCACA 318
QY 480 TACAAGACGACCTGAACCTACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTG 539
Db 319 TACACCCAGCTCTGGAACCTGGGTTCCTCGAGTGTCTGAGCTGTGGCTCCCG---CTGTAGC 375
QY 540 ATGGCCCTCGAGGAGATTGCCCCCTGACACAGCAACGAGAACGACCCAGTGTGCGGCTGCAG 599
Db 376 TGTGACCAAGTGGAAACTCAAGCCTGCACTCGGGAACAGAACCGCATCTGCACCTGCGAGG 435
QY 600 CCGGGAATGTTCTGTGC 616
Db 436 CCGGCTGTACTGCGC 452

RESULT 9
PCT-US95-05423-12
; Sequence 12, Application PC/TUS9505423
; GENERAL INFORMATION:
; APPLICANT: Jack Lile
; APPLICANT: Tadahiko Kohno
; APPLICANT: Duane Bonam
; APPLICANT: Mary S. Rosendahl
; TITLE OF INVENTION: Production of Biologically Active
; TITLE OF INVENTION: Recombinant Neurotrophic Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS

```

```

; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05423
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/266,090
; FILING DATE: 27-JUNE-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/240,122
; FILING DATE: 09-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/087,912
; FILING DATE: 06-JULY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/680,681
; FILING DATE: 04-APRIL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/594,126
; FILING DATE: 09-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/547,750
; FILING DATE: 02-JULY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/505,441
; FILING DATE: 06-APRIL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: SYNE200/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 691 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US95-05423-12

Query Match
Best Local Similarity 3.2%; Score 62.6; DB 5; Length 691;
Matches 165; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

QY 300 TGCAGCCCGCCAGCGGTGCTCCATATGCTCGGAGAACACAGACCTGCGAGGACACAGGAA 359
Db 139 TTGCCCGCCCGCCAGGTGCGCATTTACACCTACGCTCCGGAACCGGGTCTACCTGCGGCTC 198
QY 360 AAGGAATATATGAGCCCGCCAGCAGCATCTGCTGCCGTGCGCGCCGCGCAGGACCTAT 419
Db 199 AGAGAATACTATGACACAGACAGCTCAGATGTGCTGAGCAAGTGTGCTGCGCGGCGCAACAT 258
QY 420 GTCTCAGCTAAATGTAGCGCATCCGGGACACAGTTTGGCCACATGTGCGCGAGATTC 479
Db 259 GCAAAAGTCTTCTGTACCAAGACCTCGGACACCGGTGTGTGACTCTCTGTGAGGACAGCACA 318
QY 480 TACAAGACGACCTGAACCTACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTG 539
Db 319 TACACCCAGCTCTGGAACCTGGGTTCCTCGAGTGTCTGAGCTGTGGCTCCCG---CTGTAGC 375
QY 540 ATGGCCCTCGAGGAGATTGCCCCCTGACACAGCAACGAGAACGACCCAGTGTGCGGCTGCAG 599
Db 376 TGTGACCAAGTGGAAACTCAAGCCTGCACTCGGGAACAGAACCGCATCTGCACCTGCGAGG 435
QY 600 CCGGGAATGTTCTGTGC 616
Db 436 CCGGCTGTACTGCGC 452

RESULT 10
US-09-326-394-3
; Sequence 3, Application US/09326394
; Patent No. 6306820

```


Db 61 AGAGATACTATGACGACACCTCAGATGTCTGCAGCAAAATGCTCGCCGGGGCCAAACAT 120
QY 420 GTCTCAGCTAAATGAGCGCATCCGGGACACAGATTTGTGCCACATGTGCCGAGAAATCC 479
Db 121 GCAAAAGTCTTGTACCAAGACCTCGGACACCGTGTGTGACTCTCTGTGAGGACAGCACA 180
QY 480 TACACGAGCAGTGAACCTGACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTG 539
Db 181 TACACCCAGCTGTGAACCTGGTTCCTCCGAGTGTTCGAGCTGTGGTCCCG---CTGTAGC 237
QY 540 ATGGGCTCGAGGAGATGCCCCCTGCACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 599
Db 238 TCTGACCAAGTGAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 599
QY 600 CCGGGAATGTTCTGTGC 616
Db 298 CCCGCTGTGACTGCGC 314

RESULT 14

US-09-580-235-7
; Sequence 7, Application US/09580235
; Patent No. 6433158

GENERAL INFORMATION:

APPLICANT: Pettit, Dean
TITLE OF INVENTION: Site Specific Protein Modification

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Janis C Henry

STREET: 51 University

CITY: Seattle

STATE: WA

COUNTRY: US

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/580,235

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/102,530

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Henry, Janis C

REGISTRATION NUMBER: 34,347

REFERENCE/DOCKET NUMBER: 2637

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)470-4189

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 705 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: CDS

LOCATION: 1..705

US-09-580-235-7

Query Match 3.2%; Score 62.6; DB 4; Length 705;

Best Local Similarity 52.1%; Pred. No. 9.8e-06;

Matches 165; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

QY 300 TCGAGCCCGGCGGTGCTCCATATGCTCGGAGACCAAGCAAGCAAGCAAGCAAGCAAGCAAG 359

Db 1 TTGCCCCCGCCAGGTGGCAATTTACCCCTACGCCCGGAGACCCCGGGAGACATGCGGGCTC 60
QY 360 AAGGAATACTATGAGCCCCAGCAGCGCATCTGCTGCTCCCGTGCCTCCGCGGACGACCTAT 419
Db 61 AGAGATACTATGAGCAGACAGCTCAGATGTCTGCAGCAAAATGCTCGCCGGGGCCAAACAT 120
QY 420 GTCTCAGCTAAATGAGCGCATCCGGGACACAGATTTGTGCCACATGTGCCGAGAAATCC 479
Db 121 GCAAAAGTCTTGTACCAAGACCTCGGACACCGTGTGTGACTCTCTGTGAGGACAGCACA 180
QY 480 TACACGAGCAGTGAACCTGACCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTG 539
Db 181 TACACCCAGCTGTGAACCTGGTTCCTCCGAGTGTTCGAGCTGTGGTCCCG---CTGTAGC 237
QY 540 ATGGGCTCGAGGAGATGCCCCCTGCACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 599
Db 238 TCTGACCAAGTGAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 599
QY 600 CCGGGAATGTTCTGTGC 616
Db 298 CCCGCTGTGACTGCGC 314

RESULT 15

US-09-580-181-1

; Sequence 1, Application US/09580181

; Patent No. 6441136

GENERAL INFORMATION:

APPLICANT: Pettit, Dean

TITLE OF INVENTION: Site Specific Protein Modification

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Janis C Henry

STREET: 51 University

CITY: Seattle

STATE: WA

COUNTRY: US

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/580,181

FILING DATE: 26-MAY-2000

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/102,530

FILING DATE: 22-JUN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Henry, Janis C

REGISTRATION NUMBER: 34,347

REFERENCE/DOCKET NUMBER: 2637

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206)470-4189

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 705 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: CDS

LOCATION: 1..705

US-09-580-181-1

Query Match 3.2%; Score 62.6; DB 4; Length 705;

Best Local Similarity 52.1%; Pred. No. 9.8e-06;

Matches 165; Conservative 0; Mismatches 149; Indels 3; Gaps 1;

```
Qy 300 TCGCAGCCCGAGGGTGCCTCCATATGCTCGGAGAACACAGACTGTCAGGGACCGAGAA 359
Db 1 TTGCCCCCGAGGTGGCATTTACACCCCTACGCCCGGAGCCGGGAGCACATGCCGGCTC 60
Qy 360 AAGGAATACTATGAGCCCCCAGCAGCGCATCTGCTCTCCCGCTGCCCGCCAGGCACCTAT 419
Db 61 AGAGAATACTATGAGCACAGCTCAGATGTGCTGCAGCAATGCTCGCCGGCCACAT 120
Qy 420 GTCTCAGCTAATGTAGCCGCATCGGGGACACAGTTTGTGCCACATGTGCCGAGAAATCC 479
Db 121 GCAAAAGTCTTCTGTACCAAGACCTCGGACACCGCTGTGTGACTCTCTGTGAGGACAGCACA 180
Qy 480 TACACGAGCACTGGAATCTGACCATCTGCCAGCTGTGCCGCCCTGTGACCCAGTG 539
Db 181 TACACCCAGCTCTGGAATCTGCTCCCGAGTGTGCTGTGAGCTGTGGCTCCCG---CTGTAGC 237
Qy 540 ATGGCCCTCGAGGAGATTGCCCTCGACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 599
Db 238 TCTGACCAAGTGGAAACTCAAGCCTGCACTCGGGAACAGAACCCGATCTGCACCTGCAGG 297
Qy 600 CCGGGAATGTTCTGTGC 616
Db 298 CCCGCTGGTACTGCGC 314
```

Search completed: April 15, 2003, 16:28:26
Job time : 71.8651 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:04:15 ; Search time 1210.24 seconds
(without alignments)
15606.600 Million cell updates/sec

Title: US-09-917-372-2_COPY_629_1277

Perfect score: 649

Sequence: 1 cgagtggtacactcgagc.....aacctccatgccattccc 649

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb.ba:*

2: gb.htg:*

3: gb.in:*

4: gb.om:*

5: gb.ov:*

6: gb.pat:*

7: gb.ph:*

8: gb.pl:*

9: gb.pr:*

10: gb.ro:*

11: gb.sts:*

12: gb.sy:*

13: gb.un:*

14: gb.vi:*

15: em.ba:*

16: em.fun:*

17: em.hum:*

18: em.in:*

19: em.mu:*

20: em.om:*

21: em.or:*

22: em.ov:*

23: em.pat:*

24: em.ph:*

25: em.pl:*

26: em.ro:*

27: em.sts:*

28: em.un:*

29: em.vi:*

30: em.htg_hum:*

31: em.htg_inv:*

32: em.htg_other:*

33: em.htg_mus:*

34: em.htg_pln:*

35: em.htg_rod:*

36: em.htg_mam:*

37: em.htg_vrt:*

38: em.sy:*

39: em.htgo_hum:*

40: em.htgo_mus:*

41: em.htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	531	81.8	2091	9	AK095418	AK095418 Homo sapi
2	531	81.8	2136	6	AX331906	AX331906 Sequence
3	531	81.8	2136	6	AX332212	AX332212 Sequence
4	531	81.8	2136	6	AX409488	AX409488 Sequence
5	531	81.8	2136	9	HUMTNFRP	L04270 Homo sapien
6	529.4	81.6	2148	9	AK027080	AK027080 Homo sapi
7	529.4	81.6	2161	9	BC026262	BC026262 Homo sapi
8	270	41.6	1614	10	MUSLYNPHOB	L38423 Mus musculus
9	270	41.6	2076	10	MMU29173	U29173 Mus musculus
c 10	235.4	36.3	140026	9	AC005840	AC005840 Homo sapi
c 11	198	30.5	1605	9	HUMTUMNEC	L04489 Homo sapien
c 12	117.6	18.1	178228	2	AC128082	AC128082 Rattus no
c 13	99	15.3	323	9	HS27B9R	Z60529 H.sapiens C
c 14	97.6	15.0	193	9	HS27B9F	Z60528 H.sapiens C
c 15	82.4	12.7	187998	2	AC125909	AC125909 Rattus no
c 16	72	11.1	187998	2	AC125909	AC125909 Rattus no
c 17	66.8	10.3	2048	10	MMU30798	U30798 Mus musculus
c 18	45	6.9	125020	9	AF429315	AF429315 Homo sapi
c 19	41.6	6.4	125020	9	AF429315	AF429315 Homo sapi
c 20	41.4	6.4	107109	9	HSBJ781B1	AL118522 Human DNA
c 21	41	6.3	58311	2	AC110596	AC110596 Homo sapi
c 22	40.8	6.3	837	6	AX027018	AX027018 Sequence
c 23	40.8	6.3	845	6	AX027450	AX027450 Sequence
c 24	40.8	6.3	1276	4	AF248545	AF248545 Sus scrof
c 25	40	6.2	134092	2	H0702605	AL442105 Oryza sat
c 26	39.8	6.1	167303	9	AP001527	AP001527 Homo sapi
c 27	39.8	6.1	189637	2	AC009765	AC009765 Homo sapi
c 28	39.8	6.1	199891	9	CNSO0480	AL079303 Human chr
c 29	39.4	6.1	91927	9	AC004771	AC004771 Homo sapi
c 30	39.2	6.0	519	6	AR076920	AR076920 Sequence
c 31	39.2	6.0	519	6	AR078310	AR078310 Sequence
c 32	39.2	6.0	519	6	AR085413	AR085413 Sequence
c 33	39.2	6.0	519	6	AR103369	AR103369 Sequence
c 34	39.2	6.0	519	6	AR169226	AR169226 Sequence
c 35	39.2	6.0	834	6	AR027007	AR027007 Sequence
c 36	39.2	6.0	1004	6	AR138075	AR138075 Sequence
c 37	39.2	6.0	1004	6	AR153584	AR153584 Sequence
c 38	39.2	6.0	1004	6	I07284	I07284 Sequence 31
c 39	39.2	6.0	1004	9	HSCDW40	X60592 Human CD40
c 40	39.2	6.0	1157	9	BC012419	BC012419 Homo sapi
c 41	39.2	6.0	154011	2	AC018571	AC018571 Homo sapi
c 42	39.2	6.0	159983	2	AC040958	AC040958 Homo sapi
c 43	39.2	6.0	191583	9	AC015651	AC015651 Homo sapi
c 44	39.2	6.0	203613	2	AC118575	AC118575 Lemur cat
c 45	39	6.0	73565	2	AC121318	AC121318 Mus muscu

ALIGNMENTS

RESULT 1
AK095418
LOCUS
DEFINITION
Homo sapiens CDNA FLJ38099 fis, clone D30ST1000238, highly similar
to LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
ACCESSION
AK095418
VERSION
AK095418.1 GI:21754669
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens CD34+ Cells cDNA to mRNA, clone_lib:D30ST1
clone:D30ST1000238.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,

TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2091)
AUTHORS	Isogai, T. and Yamamoto, J.
TITLE	Direct Submission
JOURNAL	Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatori, Kisarazu, Chiba 297-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

```

FEATURES
  source
    i. 2091
      location/Qualifiers
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="P3OST1000238"
        /cell_type="CD34+ Cells"
        /clone_lib="P3OST1"
        /note="Cloning vector: pMEI8SFL3-mRNA from CD34+ cells
              after 3-days OD induction. -primary culture, CD34+ Cells"
BASE COUNT
  440 a 691 c 594 g 366 t

```

Query Match	81.8%;	Score 531;	DB 9;	Length 2091;
Best Local Similarity	85.7%;	Pred. No. 2.6e+129;		
Matches 649;	Conservative	0;	Mismatches 0;	Indels 108;
Gaps				1;

QY	1	CGAGTGATACACATGCGAGGCTACTTTCTGACTGCCCGCCTGGCACTGAAGCCGAGCTCAA	60
Db	566	CGAGTGATACACATGCGAGGCTACTTTCTGACTGCCCGCCTGGCACTGAAGCCGAGCTCAA	625
QY	61	AGATGAAGTTGGGAAGGTTACACCACTGGTCCCTGCAAGCGAGGGCACTTCCAGAA	120
Db	626	AGATGAAGTTGGGAAGGTTACACCACTGGTCCCTGCAAGCGAGGGCACTTCCAGAA	685
QY	121	TACCTCCTCCCCAGCGCCGGCTGCCAGCCCAACACAGGTTGTGAGAACCAAGGCTCTGGT	180
Db	686	TACCTCCTCCCCAGCGCCGGCTGCCAGCCCAACACAGGTTGTGAGAACCAAGGCTCTGGT	745
QY	181	GGAGGCACTCCAGGCACTGCCAGTCGGACACAACTGCCAAAATCATTTAGAGCCACT	240
Db	746	GGAGGCACTCCAGGCACTGCCAGTCGGACACAACTGCCAAAATCATTTAGAGCCACT	805
QY	241	GCCCCAGAGATGCA-----	256
Db	806	GCCCCAGAGATGCAAGAACCATGCTGATGCTGGCGTTCTGCTGCCACTGGCGCTCTT	865
QY	257	-----	256
Db	866	TTCTGCTCTTTGCCACCGTCTTCCTGCATCTGGAAGAGCCACCTTCTCTCTGCAGAA	925
QY	257	---GGATCGTGTCAAGAGCGTCCGCGAGGAGAGGGACCCAAATCGTGTAGCTGGAAG	312
Db	926	ACTGGGATCGTGTCAAGAGCGTCCGCGAGGAGAGGGACCCAAATCGTGTAGCTGGAAG	985
QY	313	CTGGAGCCCTCCGAAGGCCATCCATACTTCCCTGACTTGGTACAGCCACTGCTACCCAT	372
Db	986	CTGGAGCCCTCCGAAGGCCATCCATACTTCCCTGACTTGGTACAGCCACTGCTACCCAT	1045
QY	373	TTCTGGAGATGTTTCCCCAGTATCCACTGGGCTCCCGCAGCCCCAGTTTTTGGAGGCAGG	432

Db	1046	TTCTGGAGATGTTTCCCAAGTATCCACTGGGCTCCCGCAGCCCAAGTCTTTGGAGGCAGG	1105
Qy	433	GGTCCGCAACAGCAGAGTCTCTGGAGCTGACACAGGAGCGCAGTTGGAACCCGGGA	492
Db	1106	GGTCCGCAACAGCAGAGTCTCTGGAGCTGACACAGGAGCGCAGTTGGAACCCGGGA	1165
Qy	493	CGAGAGCAGGTGCCACCGGTACCAATGGCATTCATGTACCGGCGGGTCTATGACTAT	552
Db	1166	CGAGAGCAGGTGCCACCGGTACCAATGGCATTCATGTACCGGCGGGTCTATGACTAT	1225
Qy	553	CAC'TGGCAACATCTACATCTACATGACACAGTACTGGGGGACCACCGGTCTGAGAA	612
Db	1226	CAC'TGGCAACATCTACATCTACATGACACAGTACTGGGGGACCACCGGTCTGAGAA	1285
Qy	613	CCTCCAGCTACCCCGGAACCTCCATACCCGATTCCTC	649
Db	1286	CCTCCAGCTACCCCGGAACCTCCATACCCGATTCCTC	1322
RESULT 2			
AX331906			
LOCUS	AX331906	2136 bp	DNA linear PAT 09-JAN-2002
DEFINITION	Sequence 2415 from Patent WO0194629.		
ACCESSION	AX331906		
VERSION	AX331906.1	GI:18122540	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Young, P. E., Augustus M., Carter, K. C., Ebner, R., Endress, G., Horrigan, S., Soppet, D. R. and Weaver, Z.		
TITLE	Cancer gene determination and therapeutic screening using signature gene sets		
JOURNAL	Patent: WO 0194629-A 2415 13-DEC-2001;		
FEATURES	Avalon Pharmaceuticals (US)		
source	Location/Qualifiers		
	1..2136	/organism="Homo sapiens"	
	446 a	706 c	376 t
BASE COUNT			
ORIGIN			
Query Match	81.88;	Score 531;	DB 6; Length 2136;
Best Local Similarity	85.7%;	Pred. No. 2.6e-129;	
Matches 649;	Conservative 0;	Mismatches 0;	Indels 108; Gaps 1;
Qy	1	CGAGTGTACACATGCGAGCTACTTCTGACTGCGCCCTGGCACTGAAGCGAGCTCAA	60
Db	579	CGAGTGTACACATGCGAGCTACTTCTGACTGCGCCCTGGCACTGAAGCGAGCTCAA	638
Qy	61	AGATGAAGTTGGGAAGGTTAAACACACTCGTGCCTCGAAGCAGGCACTTCCAGAA	120
Db	639	AGATGAAGTTGGGAAGGTTAAACACACTCGTGCCTCGAAGCAGGCACTTCCAGAA	698
Qy	121	TACCTCTCCCCAGCGCCGCTGCCAGCCCCACACAGGTGTGAGAACCAAGTCTGGT	180
Db	699	TACCTCTCCCCAGCGCCGCTGCCAGCCCCACACAGGTGTGAGAACCAAGTCTGGT	758
Qy	181	GGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAATCCATTAGAGCCACT	240
Db	759	GGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAATCCATTAGAGCCACT	818
Qy	241	GCCCCAGAGATGTC	256
Db	819	GCCCCAGAGATGTCAGGAACCATGCTGATGCTGGCGGTTCTGCTGCCACTGGCCCTCTT	878
Qy	257	-----	256
Db	879	TCTGCTCTTGCCACCGTCTCTCTGATCTGGAAGAGCAACCTTCTCTCTGCAGAA	938
Qy	257	----GGATCGCTGCTCAAGAGCGCTCCGAGGAGAGGAGGACCAATCTCTAGCTGGAAG	312

```

Db 939 ACTGGATCGCTCTCAAGAGCGTCCGAGGAGAGGAGCCCAATCCTGTAGTGGAG 998
QY 313 CTGGAGCCTCCGAAGGCCATCCATACCTTCCCTGACTTGGTACAGCCACTGCTACCCAT 372
Db 999 CTGGAGCCTCCGAAGGCCATCCATACCTTCCCTGACTTGGTACAGCCACTGCTACCCAT 1058
QY 373 TTCTGGAGATGTTTCCCAAGTATCCACTGGGTCTCCCGCAGGCCCAAGTTTGGAGGCAGG 432
Db 1059 TTCTGGAGATGTTTCCCAAGTATCCACTGGGTCTCCCGCAGGCCCAAGTTTGGAGGCAGG 1118
QY 433 GTGGCCGAACAGCAGAGTCTCTGTGACCTGACAGGAGGCCGAGTTGGAACCCGGGGA 492
Db 1119 GTGGCCGAACAGCAGAGTCTCTGTGACCTGACAGGAGGCCGAGTTGGAACCCGGGGA 1178
QY 493 GCAGAGCAGGTGGCCACGCTACCAATGCAATGCTATGTCACCGCGGCTCTATGACTAT 552
Db 1179 GCAGAGCAGGTGGCCACGCTACCAATGCAATGCTATGTCACCGCGGCTCTATGACTAT 1238
QY 553 CACTGGCAACATCTACATCTACAATGGACAGTACTGGGGGAGCACCAACCGGGTCTTGAGA 612
Db 1239 CACTGGCAACATCTACATCTACAATGGACAGTACTGGGGGAGCACCAACCGGGTCTTGAGA 1298
QY 613 CCTCCAGCTACCCCGAACCTCCATACCCCATTTCCC 649
Db 1299 CCTCCAGCTACCCCGAACCTCCATACCCCATTTCCC 1335

RESULT 3
AX332212 AX332212 2136 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 2721 from Patent WO0194629.
DEFINITION AX332212
ACCESSION AX332212
VERSION AX332212.1 GI:18122846
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrihan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 2721 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
Location/Qualifiers
1..2136
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 446 a 706 c 608 g 376 t
ORIGIN
Query Match 81.8%; Score 531; DB 6; Length 2136;
Best Local Similarity 85.7%; Pred. No. 2.6e-129;
Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

QY 1 CGAGTGTACACTGCGAGCTACTTCTGACTGCGCGCTGGCAGCTGAAGCCGAGTCAA 60
Db 579 CGAGTGTACACTGCGAGCTACTTCTGACTGCGCGCTGGCAGCTGAAGCCGAGTCAA 638
QY 61 AGATGAAGTTGGAGGGTAACAACCACTGCGTCCCTCGAAGCAGGCACCTTCCAGAA 120
Db 639 AGATGAAGTTGGAGGGTAACAACCACTGCGTCCCTCGAAGCAGGCACCTTCCAGAA 698
QY 121 TACCTCTCCCGAGCCCGCTGCGAGCCGCCACACAGGTGTGAGAACCAAGTCTGTT 180
Db 699 TACCTCTCCCGAGCCCGCTGCGAGCCGCCACACAGGTGTGAGAACCAAGTCTGTT 758
QY 181 GGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAAATTCATTTAGAGCCAT 240
Db 759 GGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAAATTCATTTAGAGCCAT 818

```

```

QY 241 GCCCCAGAGATGTCA----- 256
Db 819 GCCCCAGAGATGTCAAGAACCATGCTGATGCTGGCGTCTCTGCTGCCACTGSCCTTCTT 878
QY 257 ----- 256
Db 879 TCTGCTCCTTGGCAGCGTCTTCTCTGCACTTGGAGAGCCACCTTCTCTCTGCAAGAA 938
QY 257 ----GGATCGCTCTCAAGAGCGTCCGAGGAGAGGAGGCCCAATCCTGTAGTGGAG 312
Db 939 ACTGGAGTCCGCTCTCAAGAGCGTCCGAGGAGAGGAGGCCCAATCCTGTAGTGGAG 998
QY 313 CTGGAGCCTCCGAAGGCCATCCATACCTTCCCTGACTTGGTACAGCCACTGCTACCCAT 372
Db 999 CTGGAGCCTCCGAAGGCCATCCATACCTTCCCTGACTTGGTACAGCCACTGCTACCCAT 1058
QY 373 TTCTGGAGATGTTTCCCAAGTATCCACTGGGTCTCCCGCAGGCCCAAGTTTGGAGGCAGG 432
Db 1059 TTCTGGAGATGTTTCCCAAGTATCCACTGGGTCTCCCGCAGGCCCAAGTTTGGAGGCAGG 1118
QY 433 GTGGCCGAACAGCAGAGTCTCTGTGACCTGACAGGAGGCCGAGTTGGAACCCGGGGA 492
Db 1119 GTGGCCGAACAGCAGAGTCTCTGTGACCTGACAGGAGGCCGAGTTGGAACCCGGGGA 1178
QY 493 GCAGAGCAGGTGGCCACGCTACCAATGGCAATGCTATGTCACCGCGGCTCTATGACTAT 552
Db 1179 GCAGAGCAGGTGGCCACGCTACCAATGGCAATGCTATGTCACCGCGGCTCTATGACTAT 1238
QY 553 CACTGGCAACATCTACATCTACAATGGACAGTACTGGGGGAGCACCAACCGGGTCTTGAGA 612
Db 1239 CACTGGCAACATCTACATCTACAATGGACAGTACTGGGGGAGCACCAACCGGGTCTTGAGA 1298
QY 613 CTCCCGAGTACCCCGAACCTCCATACCCCATTTCCC 649
Db 1299 CTCCCGAGTACCCCGAACCTCCATACCCCATTTCCC 1335

RESULT 4
AX409488 AX409488 2136 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 2135 from Patent WO0229103.
DEFINITION AX409488
ACCESSION AX409488
VERSION AX409488.1 GI:21442193
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2135 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source
Location/Qualifiers
1..2136
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 446 a 706 c 608 g 376 t
ORIGIN
Query Match 81.8%; Score 531; DB 6; Length 2136;
Best Local Similarity 85.7%; Pred. No. 2.6e-129;
Matches 649; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

QY 1 CGAGTGTACACTGCGAGCTACTTCTGACTGCGCGCTGGCAGCTGAAGCCGAGTCAA 60
Db 579 CGAGTGTACACTGCGAGCTACTTCTGACTGCGCGCTGGCAGCTGAAGCCGAGTCAA 638
QY 61 AGATGAAGTTGGAGGGTAACAACCACTGCGTCCCTCGAAGCAGGCACCTTCCAGAA 120
Db 639 AGATGAAGTTGGAGGGTAACAACCACTGCGTCCCTCGAAGCAGGCACCTTCCAGAA 698

```


FEATURES

source	Location/Qualifiers
1. .2161	

```
/organism="Homo sapiens"
/db_xref="LocusID:4055"
/db_xref="taxon:9606"
/clone="MGC:22593 IMAGE:4703362"
/tissue_type="lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
199..1506
/codon_start=1
/product="lymphotoxin beta receptor (TNFR superfamily, member 3)"
/protein_id="AAH26262.1"
/db_xref="GI:20072213"
/translation="MLPWTAPGLAWPLVLGLLAAASQAPQAPVYASNOTCR
DDEKEYEPOHRIICCSRCPPGYVSASKSRIQVTCATCAENSYNHWNLTQCLCR
PCDVMGLIEAPCTSKRTQRCQFCAAWALECTHCELLSDCPGTAEALKDEY
KGNNHVECKAGHFTSPSARCQPHTRCENQGLVEAAPGTADQCTKKNPLELP
PENSHTMLMLAVLLPLAFFLLATVFCISWKSRLKGLSKLRPPQGGEPNPVAG
SWEPKHAHFFDLPVQLPPLISGDSVSTGLPAAPVLEAGVQQQSPDLDTREPOLE
PGQSOVAGHTNGIHVTGGSMTITGNIYIYNGPVLGGPPGDLPATPEPPYPIPEG
DPGPPGLSTPHQDGRKAWHLAETEHGATPSNRGPRNQFIHID"
BASE COUNT 473 a 708 c 606 g 374 t
ORIGIN
Query Match 81.6%; Score 529.4; DB 9; Length 2161;
Best Local Similarity 85.6%; Pred. No. 7e-129;
Matches 648; Conservative 0; Mismatches 1; Indels 108; Gaps 1;
QY 1 CGAGTGATACACTGCGAGCTACTTTCTGACTGCGCCGCTGGCACTGAAGCCGAGCTCAA 60
Db 609 CGAGTGATACACTGCGAGCTACTTTCTGACTGCGCCGCTGGCACTGAAGCCGAGCTCAA 668
QY 61 AGATGAAGTTGGGAAGGTAAACACCACTGCGTCCCTGCAAGCGAGCGCACTTCCAGAA 120
Db 669 AGATGAAGTTGGGAAGGTAAACACCACTGCGTCCCTGCAAGCGCGGCACTTCCAGAA 728
QY 121 TACCTCTCTCCCGAGCGCGCTGCCAGCCCCACACCAAGGTGTGAGAACCAAGTCTGGT 180
Db 729 TACCTCTCTCCCGAGCGCGCTGCCAGCCCCACACCAAGGTGTGAGAACCAAGTCTGGT 788
QY 181 GGAGCGAGCTCCAGGCACTGCCAGTCCGACACCAACCTGCAAAAATCCATTAGAGCCACT 240
Db 789 GGAGCGAGCTCCAGGCACTGCCAGTCCGACACCAACCTGCAAAAATCCATTAGAGCCACT 848
QY 241 GCCCCAGAGATGCA----- 256
Db 849 GCCCCAGAGATGTCAGGAACCACTGTGATGCTGCGCGCTTCTGCTGCCACTGGCCTTCTT 908
QY 257 ----- 256
Db 909 TCTGCTCTTGGCCACCGTCTTCTCTGATCTGGAAGAGCCACCCCTTCTCTGTCAGGAA 968
QY 257 ----GGATCGTGTCAAGAGGCGTCCGAGGAGAGAGGAGCCCAATCTGTAGTGAAG 312
Db 969 ACTGGGATCGTGTCAAGAGGCGTCCGAGGAGAGGAGGAGCCCAATCTGTAGTGAAG 1028
QY 313 CTGGAGCCTCCGAAGCGCCATCATACTTCCCTGACTTGTGACAGCACTGCTACCCAT 372
Db 1029 CTGGAGCCTCCGAAGCGCCATCATACTTCCCTGACTTGTGACAGCACTGCTACCCAT 1088
QY 373 TTCTGGAGATGTTTTCCTCCAGTATCCACTGGGCTCCCGAGCGCCAGTTTGGAGGAGG 432
Db 1089 TTCTGGAGATGTTTTCCTCCAGTATCCACTGGGCTCCCGAGCGCCAGTTTGGAGGAGG 1148
QY 433 GGTCCCGACAGCAGAGTCTCTGAGACTGACACAGAGGAGCGGAGTTGGAACCCGGGGA 492
Db 1149 GGTCCCGACAGCAGAGTCTCTGAGACTGACACAGGAGGAGCGGAGTTGGAACCCGGGGA 1208
QY 493 GCAGAGCAGGTGGCCACAGTACCAATGGCATTCATGTCACCGCGGCTATGACTAT 552
Db 1209 GCAGAGCAGGTGGCCACAGTACCAATGGCATTCATGTCACCGCGGCTATGACTAT 1268
```

```
QY 553 CACTGGCAACATCTACATCTACATGACACGACTGCGGGGACACCGCGGTCTCTGGAGA 612
Db 1269 CACTGGCAACATCTACATCTACATGACACGACTGCGGGGACACCGCGGTCTCTGGAGA 1328
QY 613 CACTCCAGCTACCCCGAGCTCCATACCCATCC 649
Db 1329 CACTCCAGCTACCCCGAGCTCCATACCCATCC 1365
RESULT 8
MUSLYMPHOB 1614 bp mRNA linear ROD 24-MAY-1996
LOCUS Mus musculus lymphotoxin-beta receptor gene, complete cds.
DEFINITION L38423
VERSION L38423.1 GI:600222
KEYWORDS lymphotoxin-beta receptor; transmembrane protein.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1614)
AUTHORS Nakamura,T., Tashiro,K., Nazarea,M., Nakano,T., Sasayama,S. and
Honjo,T.
TITLE The murine lymphotoxin-beta receptor cDNA: isolation by the signal
sequence trap and chromosomal mapping
JOURNAL Genomics 30 (2), 312-319 (1995)
MEDLINE 96163885
PUBMED 8586432
FEATURES
source
1..1614
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="ST-2"
186..1433
/note="putative"
/codon_start=1
/product="lymphotoxin-beta receptor"
/db_protein_id="AAB00846.1"
/db_xref="GI:600223"
/translation="MRLPRASPCGLAWPLLLGLSLVLSQPLVPPRYENOTCM
DDEKEYEPMHDVCCSRCPGFVAVCSRODTCKTCKPHNSYHWNHLSLSTCLCR
PCDVLGFEEVAFCTSDRAECRCQPMSCVYLDNECVHERLVLCOPTAEAVTD
EIMDITNCVCPKPHQNTSSPRACQPHTRCETIGLVEAFPGYSYDITCKNPEP
GAMLLAILLSLVLLFTTFLACAMRHPSLCRKLGLTKLRHPEGESEPPCPAPRAD
PHFDLAEPPLPMSGLSPSPAGPTAPSLVEEVQQQSLVQARLEAEPEHGGVA
HGANGIHVTGGSVTVTNIYIYNGPVLGTRGPGDPPADPEPPYPTPEGAGPSPSLS
"TPYQEDGKAHLAETETLCCQDI."
BASE COUNT 316 a 537 c 455 g 306 t
ORIGIN
Query Match 41.6%; Score 270; DB 10; Length 1614;
Best Local Similarity 65.8%; Pred. No. 1.9e-60;
Matches 467; Conservative 0; Mismatches 165; Indels 78; Gaps 2;
QY 17 GAGTACTTTCTGACTGCGCGCTGGCACTGAAGCCGAGCTCAAAGTGAAGTTGGGAAG 76
Db 618 GAGCGGCTTGTACTCTGCCAGCTGGCACAGAAGCCGAGCTCAAGATTAATATGAT 677
QY 77 GGTAAACAACACCTGCGTCCCTCAAGGAGGCACTTCCAGATACTCTCCCTCCAGC 136
Db 678 ACTGAGCTCAACTGTGTCCCTGTAAAGCCGGGACACTTCCAGAACACTTCTCCCTCCGA 737
QY 137 GCCCGCTGCCAGCCCCACACAGGTGTGAGAACCAAGGTCTGTGGAGGAGCTCCAGGC 196
Db 738 GCCCGCTGTCAACCCCATACAGATGTGAGATCCAGGCGCTGGTGGAGGAGCTCCAGGT 797
QY 197 ACTGCGCAGTCCGACACACACTGCAAAATTCATTAGAGCACTGCCCCCAGAGATGTCA 256
Db 798 ACCTCTCTACTCGGATACCATCTGTAAATAATCCCCAGAGCCAGGACCAATGCTACTGCT 857
QY 257 GGATCGCTGCTCAAG-----AGGCGTCCGAGGG 285
Db 858 GCCATCTGCTGCTGCTGGTGGTCTTCTTCTCTTACCACACTGCTGTCGCTGTGCTGCTGG 917
```


TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

Direct Submission
Submitted (22-OCT-1998) Department of Molecular Biology, Albert
Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY
10461, USA
3 (bases 1 to 140026)
Montgomery, K.T., Lau, S.T. and Kucheralapati, R.
Direct Submission
Submitted (20-OCT-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY
10461, USA
On Oct 20, 2000 this sequence version replaced gi:3779004.
-----Genome Center:
Albert Einstein College of Medicine
Code: AECOM
Web site: <http://sequence.aecom.yu.edu/chr12/>
Contact: ihan@sequence.aecom.yu.edu

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human sequences.

Genes and Regions of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3399-3402) similarity (expect < 1e-34) to EST and cDNA sequences in Unigene. Genes demonstrate at least two exons flanked by consensus splice sites that maintain sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Attempts are made to complete double stranded sequence for all regions. All sequence is completed to a standard of coverage with a minimum of 3 reads with no ambiguities. If the sequence coverage for a region does not meet this standard, it is indicated in the annotation as Low Coverage. Low coverage linkages are verified by PCR product size verification or verification of forward and reverse reads from clones which span the low coverage area.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated average error rate is less than 1 per 10,000 bases using the Consed quality parameters. Regions that do not meet this requirement are annotated as Low Quality.

```

-----Summary Statistics
Center project name: RP1-102H24
Sequencing vector: M13
Chemistry: Dye-terminator Big Dye; 100%
Assembly program: Phrap version 0.990319
Contig length: 140026
Fraction of Phrap value < 40: 0.0426
Error Rate in Consed: 0.56 per 10,000 b
Number of N's in consensus: 1

```

----- Distribution of Quality < 40 Bases:

[illegible]

FEATURES	source	Location/Qualifiers
	repeat_region	1..140026
	repeat_region	/organism="Homo sapiens"
	repeat_region	/db_xref="taxon:9606"
	repeat_region	/chromosome="12"
	repeat_region	/clone="RP1-102E24"
	repeat_region	1..51
	repeat_region	/rpt_family="AluJb"
	repeat_region	52..76
	repeat_region	/rpt_family=""(CAAA)n"
	repeat_region	complement(254..367)
	repeat_region	/rpt_family="LIMB8"
	repeat_region	complement(377..649)
	repeat_region	/rpt_family="LIMB8"
	repeat_region	complement(650..941)
	repeat_region	/rpt_family="AluSc"
	repeat_region	complement(983..1090)
	repeat_region	/rpt_family="LIMB8"
	repeat_region	complement(1110..1339)
	repeat_region	/rpt_family="AluSx"
	repeat_region	complement(1341..1354)
	repeat_region	/rpt_family="LIMB8"
	repeat_region	complement(1355..1652)
	repeat_region	/rpt_family="AluSx"
	repeat_region	complement(1653..1787)
	repeat_region	/rpt_family="LIMB8"
	repeat_region	complement(1797..2090)
	repeat_region	/rpt_family="AluJo"
	repeat_region	complement(2241..2533)
	repeat_region	/rpt_family="AluSg"
	repeat_region	3376..3664
	repeat_region	/rpt_family="AluSg"
	repeat_region	3825..3850
	repeat_region	/rpt_family="AT-rich"
	gene	complement(join(4534..48517..5666..5694..5849,5889..7029))
	repeat_region	/gene="Plakophilin 2 (PKP7134..7432
	repeat_region	/rpt_family="AluSc"
	repeat_region	complement(7442..7588)
	repeat_region	/rpt_family="LIM4"
	repeat_region	8060..8245
	repeat_region	/rpt_family="LIMED"
	repeat_region	8246..8360
	repeat_region	/rpt_family="HAL1"
	repeat_region	8399..8563
	repeat_region	/rpt_family="AluSg"
	repeat_region	8564..9933
	repeat_region	/rpt_family="SVA"
	repeat_region	9934..10088
	repeat_region	/rpt_family="LIMD2"
	repeat_region	10107..10188
	repeat_region	/rpt_family="FLAM_A"
	repeat_region	10194..10975
	repeat_region	/rpt_family="LIMD2"
	repeat_region	10976..11283
	repeat_region	/rpt_family="AluY"
	repeat_region	complement(11303..11589)
	repeat_region	/rpt_family="AluJb"
	repeat_region	complement(11678..11692)
	repeat_region	/rpt_family="FRAM"
	repeat_region	complement(11693..11995)
	repeat_region	/rpt_family="AluSx"
	repeat_region	complement(11996..12157)
	repeat_region	/rpt_family="FRAM"
	repeat_region	12547..12899
	repeat_region	/rpt_family="MER104"
	repeat_region	complement(12900..13198)
	repeat_region	/rpt_family="AluSx"
	repeat_region	13199..13254
	repeat_region	/rpt_family="MER104"


```

repeat_region complement(13309..13472)
STS /rpt_family="L2"
13793..14017
/standard_name="GDB:384731"
repeat_region complement(16498..17181)
/rpt_family="THE1A-int"
repeat_region complement(17193..17563)
/rpt_family="THE1A-int"
repeat_region complement(17564..17915)
/rpt_family="THE1A"
repeat_region complement(18210..18342)
/rpt_family="FLAM-A"
gene join(18391..18519,18744..18902,19803..19853)
/feature="Vesicle-associated membrane protein 1 (SVB1)"
19248..19275
/rpt_family="CA)n"
repeat_region 19276..19351
/rpt_family="L1MED"
repeat_region complement(19296..19475)
/rpt_family="L1MCA"
repeat_region 19914..19933
/rpt_family="(CTG)n"
STS 20287..20651
/standard_name="GDB:214816"
STS 20290..20527
/standard_name="sts-M36200"
STS 20459..20650
/standard_name="sts-M36200"
STS 20841..21018
/standard_name="SHGC-557"
STS 20971..21045
/standard_name="D12S111E"
STS 22216..22330
/standard_name="Cda1eb01"
STS 22518..22670
/standard_name="stS4862"
gene complement(join(22595..22710,23806..23889,25797..26099,
27000..27338,31028..31297,31447..31677,32432..32495))
Query Match 36.3%; Score 235.4; DB 9; Length 140026;
Best Local Similarity 99.6%; Pred. No. 2.3e-51;
Matches 236; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 278 CGCAGGAGGAGGACCACTCTAGCTGGAGCTGGAGCCTCCGAAGGCCCATCCA 337
|||||
Db 94638 CTGACGGAGAGGAGGACCACTCTAGCTGGAGCTGGAGCCTCCGAAGGCCCATCCA 94579

QY 338 TACTTCCCTGACTGGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTATCC 397
|||||
Db 94578 TACTTCCCTGACTGGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTATCC 94519

QY 398 ACTGGGCTCCCGCAGCCCGCAGTTTGGAGGAGGAGGTCGCAACAGCAGAGTCTCTG 457
|||||
Db 94518 ACTGGGCTCCCGCAGCCCGCAGTTTGGAGGAGGAGGTCGCAACAGCAGAGTCTCTG 94459

QY 458 GACCTGACAGGAGCCCGCAGTTGGAACCGGGGAGCAGAGCAGGTGCCACCGT 514
|||||
Db 94458 GACCTGACAGGAGCCCGCAGTTGGAACCGGGGAGCAGAGCAGGTGCCACCGT 94402

RESULT 11
LOCUS HUMTUMNEC 1605 bp mRNA linear PRI 14-JAN-1995
DEFINITION Homo sapiens (clone NC18) tumor necrosis factor receptor related
protein mRNA, complete exon and repeat region.
ACCESSION L04489
VERSION L04489.1 GI:340022
KEYWORDS tumor necrosis factor receptor related protein.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1605)

```

```

AUTHORS Baens,M., Chaffanet,M., Cassiman,J.J., van den Berghe,H. and
Marynen,P.
TITLE Construction and evaluation of a hncDNA library of human 12p
transcribed sequences derived from a somatic cell hybrid
JOURNAL Genomics 16 (1), 214-218 (1993)
MEDLINE 93252381
PUBMED 8486360
FEATURES
Location/Qualifiers
1..1605
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="12p13"
/cell_line="M28 somatic cell hybrid"
/tissue_lib="hncdna tc651"
1..1605
/gene="TNFR"
repeat_region 1..105
/partial
/feature="putative"
/rpt_family="Alu"
/rpt_type-dispersed
250..444
/gene="TNFR"
/exon
/feature="L04270 sequence homologue"
/feature="experimental"
BASE COUNT 472 a 394 c 323 g 416 t
ORIGIN

Query Match 30.5%; Score 198; DB 9; Length 1605;
Best Local Similarity 100.0%; Pred. No. 1.8e-41;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 AGATCAAGTTGGAGGTTAAACAACCACTGCTGCTCCCTGCAAGCAGGCGCTCCAGAA 120
|||||
Db 248 AGATCAAGTTGGAGGTTAAACAACCACTGCTGCTCCCTGCAAGCAGGCGCTCCAGAA 307

QY 121 TACCTCTCTCCCGCAGCGCGCTGCCAGCCGACACACAGGTGTGAGAACCAAGTCTGT 180
|||||
Db 308 TACCTCTCTCCCGCAGCGCGCTGCCAGCCGACACACAGGTGTGAGAACCAAGTCTGT 367

QY 181 GGAGCAGCTCCAGCAGTCCCGCAGTCCGAGCAGCAGCAGTCCGAGCAGTCCAGTCC 240
|||||
Db 368 GGAGCAGCTCCAGCAGTCCCGCAGTCCGAGCAGCAGTCCGAGCAGTCCAGTCCAGTCC 427

QY 241 GCCCCAGAGATGTCCAGG 258
|||||
Db 428 GCCCCAGAGATGTCCAGG 445

RESULT 12
LOCUS AC128082/c 178228 bp DNA linear HTG 19-JUL-2002
DEFINITION Rattus norvegicus clone CH230-362C16, *** SEQUENCING IN PROGRESS
ACCESSION AC128082
VERSION AC128082.1 GI:21908679
KEYWORDS HTG; HTGS; PHASEL.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 178228)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,J.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,F.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burck,p., Burrell,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleaveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

```

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, J., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
 Homisi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, E., Lichtarge, O., Lieu, C., Lucier, R., Luna, R., Ma, J.,
 Lozadó, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
 Sodergren, E., Sonaik, I., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 2 (bases 1 to 178228)
 Worley K.C.
 Direct Submission
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

 Center: Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: GZQW
 Center clone name: CH230-362C16

 Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 116771 bases at least Q40
 Consensus quality: 123909 bases at least Q30
 Consensus quality: 129462 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 64 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1173: contig of 1173 bp in length
 * 1174 1273: gap of unknown length
 * 1274 2435: contig of 1162 bp in length
 * 2436 2535: gap of unknown length
 * 2535 3567: contig of 1032 bp in length
 * 3568 3667: gap of unknown length
 * 3668 5230: contig of 1563 bp in length
 * 5231 5330: gap of unknown length
 * 5331 6732: contig of 1402 bp in length

6733 6832: gap of unknown length
 6833 7990: contig of 1158 bp in length
 7991 8090: gap of unknown length
 8091 9641: contig of 1551 bp in length
 9642 9741: gap of unknown length
 9742 10776: contig of 1035 bp in length
 10777 10876: gap of unknown length
 10877 11905: contig of 1029 bp in length
 11906 12005: gap of unknown length
 12006 13200: contig of 1315 bp in length
 13201 13420: gap of unknown length
 13421 14778: contig of 1258 bp in length
 14779 14778: gap of unknown length
 14780 15851: contig of 1073 bp in length
 15852 15951: gap of unknown length
 15952 17593: contig of 1642 bp in length
 17594 17993: gap of unknown length
 17994 18911: contig of 1218 bp in length
 18912 19011: gap of unknown length
 19012 20145: contig of 1134 bp in length
 20146 20245: gap of unknown length
 20246 21556: contig of 1311 bp in length
 21557 21556: gap of unknown length
 21558 22715: contig of 1059 bp in length
 22716 22815: gap of unknown length
 22816 23949: contig of 1134 bp in length
 23950 24049: gap of unknown length
 24050 25331: contig of 1282 bp in length
 25332 25431: gap of unknown length
 25432 26617: contig of 1186 bp in length
 26618 26717: gap of unknown length
 26718 28435: contig of 1718 bp in length
 28436 28535: gap of unknown length
 28536 30660: contig of 2125 bp in length
 30661 30760: gap of unknown length
 30761 32263: contig of 1503 bp in length
 32264 32363: gap of unknown length
 32365 34165: contig of 1802 bp in length
 34166 34265: gap of unknown length
 34266 35302: contig of 1037 bp in length
 35303 35402: gap of unknown length
 35403 37282: contig of 1880 bp in length
 37283 37382: gap of unknown length
 37383 39260: gap of unknown length
 39261 40880: contig of 1620 bp in length
 40881 40980: gap of unknown length
 40982 42436: contig of 1456 bp in length
 42437 42536: gap of unknown length
 42537 44381: contig of 1845 bp in length
 44382 44481: gap of unknown length
 44482 46391: contig of 1910 bp in length
 46392 46491: gap of unknown length
 46492 48409: contig of 1918 bp in length
 48410 48509: gap of unknown length
 48510 49931: contig of 1422 bp in length
 49932 50032: gap of unknown length
 50033 51372: contig of 1340 bp in length
 51373 51471: gap of unknown length
 51472 53362: contig of 1891 bp in length
 53363 53462: gap of unknown length
 53463 55270: contig of 1708 bp in length
 55271 56697: contig of 1427 bp in length
 56698 56797: gap of unknown length
 56798 58037: contig of 1240 bp in length
 58038 58137: gap of unknown length
 58138 60361: contig of 2224 bp in length
 60362 62335: contig of 1874 bp in length
 62336 62435: gap of unknown length
 62436 64906: contig of 2471 bp in length
 64907 65006: gap of unknown length

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

```
* 65007 67475: contig of 2469 bp in length
* 67476 67575: gap of unknown length
* 67576 69653: contig of 2078 bp in length
* 69654 69753: gap of unknown length
* 69754 73830: contig of 4077 bp in length
* 73831 73930: gap of unknown length
* 73931 77047: contig of 3117 bp in length
* 77048 77147: gap of unknown length
* 77148 80438: contig of 3291 bp in length
* 80439 80538: gap of unknown length
* 80539 83532: contig of 2994 bp in length
* 83533 83632: gap of unknown length
* 83633 85754: contig of 2122 bp in length
* 85755 85854: gap of unknown length
* 85855 89203: contig of 3349 bp in length
* 89204 89303: gap of unknown length
* 89304 92865: contig of 3562 bp in length
* 92866 92965: gap of unknown length
* 92966 95763: contig of 2798 bp in length
* 95764 95863: gap of unknown length
* 95864 99798: contig of 3835 bp in length
* 99799 104352: contig of 4554 bp in length
* 104353 104453: gap of unknown length
* 104453 108165: contig of 3713 bp in length
* 108166 108265: gap of unknown length
* 111946 112045: gap of unknown length
* 112046 116301: contig of 4256 bp in length
* 116302 116401: gap of unknown length
* 116402 120825: contig of 4424 bp in length
* 120826 120925: gap of unknown length

Query Match 18.1%; Score 117.6; DB 2; Length 178228;
Best Local Similarity 72.2%; Pred. No. 2.7e-20;
Matches 153; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 313 CTGGAGCCTCGAAGGCCCATCTACTTCCTCGTACAGCCACTGCTACCCAT 372
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152659 CTGCCAGCCTCAAGAGCCACACATTCCTCGTACCTGGCAGCCACTCTACCCAT 152600

QY 373 TTTCTGAGATGTTCCCAAGTATCCACTGGGCTCCCGCAGCCAGTCTTTGGAGCAGG 432
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152599 GTCTGGAGACTTGTCCCAAGCCCTCTGTGAGCCGCCCTCTCTGGAGGAGT 152540

QY 433 GTGGCCGCAACAGCAGAGCTCTGACCTGACAGGAGCGCAGTGTGAACCCGGGCA 492
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152539 GTGTCTACAGCAGCATAGTCCCTGTATCCAGCCAGGAGCTGGAGCTGGCGGA 152480

QY 493 GCAGAGCCAGGTGGCCCGACGCTACCAATGGCA 524
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152479 ACATGCCAGGTGGCCCGACGCTGAGCTGGGCA 152448

RESULT 13
HS27B9R
LOCUS HS27B9R 323 bp DNA linear PRI 19-OCT-1995
DEFINITION H.sapiens CpG island DNA genomic Msel fragment, clone 27B9, reverse
read cp927b9.rftla.
ACCESSION Z60529
VERSION Z60529.1 GI:1032633
KEYWORDS CpG island; genomic Msel fragment.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 323)
AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk
REFERENCE 2 (bases 1 to 323)
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
Purification of CpG islands using a methylated DNA binding column
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.
```

```

Purification of CpG islands using a methylated DNA binding column
Nat. Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
PUBMED 8012384
COMMENT Vector: pGEM-52f(-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES
source
1..323
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="27b9"
/sex="male"
/tissue_type="blood"
/clone_lib="CGI-1"
/dev_stage="adult"
BASE COUNT 64 a 123 c 75 g 56 t 5 others
ORIGIN
Query Match 15.3%; Score 99; DB 9; Length 323;
Best Local Similarity 92.7%; Pred. No. 2.5e-15;
Matches 102; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 61 AGATGAAGTTGGAGGGTAACAACACCTGCGTCCCTGCAAGGCGGCACTTCAGAA 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 75 AGATGAAGTTGGAGGGTAACAACACCTGCGTCCCTGCAAGGCGGCACTTCAGAA 134

QY 121 TACCTCTCCCGCAGCGCCGCTGCAGCCGCCACACAGTGTGAGAACC 170
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 135 TACCTCTCCCGCAGCGCCGCTGCAGCCGCCACACAGTGTGAGTGCAGC 184

RESULT 14
HS27B9F/c
LOCUS HS27B9F 193 bp DNA linear PRI 19-OCT-1995
DEFINITION H.sapiens CpG island DNA genomic Msel fragment, clone 27B9, forward
read cp927b9.rftla.
ACCESSION Z60528
VERSION Z60528.1 GI:1032632
KEYWORDS CpG island; genomic Msel fragment.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 193)
AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk
REFERENCE 2 (bases 1 to 193)
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
Purification of CpG islands using a methylated DNA binding column
Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
http://www.hgmp.mrc.ac.uk/ for details
or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES
source
1..193
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="27b9"
/sex="male"
/tissue_type="blood"
/clone_lib="CGI-1"
/dev_stage="adult"
BASE COUNT 36 a 60 c 48 g 44 t 5 others
ORIGIN
```


* 29158	30973:	contig of 1816 bp in length
* 30974	31073:	gap of unknown length
* 32838	32838:	contig of 1765 bp in length
* 32839	32838:	gap of unknown length
* 32939	34616:	contig of 1678 bp in length
* 34617	34716:	gap of unknown length
* 34717	37198:	contig of 2482 bp in length
* 37199	37298:	gap of unknown length
* 37299	39477:	contig of 2179 bp in length
* 39478	39577:	gap of unknown length
* 39578	41815:	contig of 2238 bp in length
* 41816	41915:	gap of unknown length
* 41916	44488:	contig of 2573 bp in length
* 44489	44588:	gap of unknown length
* 44589	46062:	contig of 1474 bp in length
* 46063	46162:	gap of unknown length
* 46163	48328:	contig of 2166 bp in length
* 48329	48428:	gap of unknown length
* 48429	51395:	contig of 2967 bp in length
* 51396	51495:	gap of unknown length
* 51496	51459:	contig of 2664 bp in length
* 54160	54259:	gap of unknown length
* 54260	56383:	contig of 2124 bp in length
* 56384	56483:	gap of unknown length
* 56484	58779:	contig of 2296 bp in length
* 58780	58879:	gap of unknown length
* 58880	60911:	contig of 2032 bp in length
* 60912	61011:	gap of unknown length
* 61012	63907:	contig of 2896 bp in length
* 63908	64007:	gap of unknown length
* 64008	66202:	contig of 2195 bp in length
* 66203	66302:	gap of unknown length
* 66303	69738:	contig of 3436 bp in length
* 69739	69838:	gap of unknown length
* 69839	72410:	contig of 2572 bp in length
* 72411	72510:	gap of unknown length
* 72511	76219:	contig of 3709 bp in length
* 76220	76319:	gap of unknown length
* 76320	79029:	contig of 2710 bp in length
* 79030	79129:	gap of unknown length
* 79130	83311:	contig of 4182 bp in length
* 83312	83411:	gap of unknown length
* 83412	86394:	contig of 2983 bp in length
* 86395	86494:	gap of unknown length
* 86495	88582:	contig of 2088 bp in length
* 88583	88682:	gap of unknown length
* 88683	90972:	contig of 2290 bp in length
* 90973	91072:	gap of unknown length
* 91073	93875:	contig of 2803 bp in length
* 93876	93975:	gap of unknown length
* 93976	98771:	contig of 4796 bp in length
* 98772	98871:	gap of unknown length
* 98872	101999:	contig of 3128 bp in length
* 102000	102099:	gap of unknown length
* 102100	105352:	contig of 3253 bp in length
* 105353	105452:	gap of unknown length
* 105453	109873:	contig of 4421 bp in length
* 109874	109873:	gap of unknown length
* 109974	112881:	contig of 2908 bp in length
* 112882	112981:	gap of unknown length
* 112982	117279:	contig of 4298 bp in length
* 117280	117379:	gap of unknown length
* 117380	121634:	contig of 4255 bp in length
* 121635	121734:	gap of unknown length
* 121735	125920:	contig of 4186 bp in length
* 125921	126020:	gap of unknown length
* 126021	129909:	contig of 3889 bp in length
* 129910	130009:	gap of unknown length
* 130010	135723:	contig of 5714 bp in length
* 135724	135823:	gap of unknown length

Query Match
Best Local Similarity

12.7%; Score 82.4; DB 2; Length 187998;
81.9%; Pred. No. 5.1e-11;

Matches	95;	Conservative	0;	Mismatches	21;	Indels	0;	Gaps	0;
QY	534	CCGGCGGGTCTATGACTATCATCTGCAACATCTACATCTACATCTACATCTGACACGACTGGGG	593						
Db	154396	CTGCAGGCTCTGTGACTGTACCCGGCAATATCTACATATACATATGAGCCAGTCTCTGGGG	154455						
QY	594	GACCACCGGGTCTTGAGACCTCCAGCTACCCCGAACCTCCATACCCCATTTCCC	649						
Db	154456	GAACACGGGGCCCTGGAGACCTCCAGCTCCCGCTGAGCCTCCATACCCGACTCCC	154511						

Search completed: April 15, 2003, 15:16:30
Job time : 1421.24 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:05:15 ; Search time 100.083 Seconds
(without alignments)
12938.300 Million cell updates/sec

Title: US-09-917-372-2_COPY_804_1378
Perfect score: 575
Sequence: 1 ctgtgtgagcgagctccagg.....cactgtgtggtgccacaccctc 575

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	85.9	1594	ABQ55003	Human ovarian anti
2	494	85.9	2136	ABN95637	Gene #2135 used to
3	494	85.9	2136	ABK64502	Human benign prost
4	494	85.9	2136	ABL64078	Breast cancer rela
5	494	85.9	2136	ABL64384	Stomach cancer rel
6	253.8	44.1	281	ABL82623	Human ovarian can
7	251.2	43.7	289	ABL82688	Human ovarian can
8	61.2	10.6	481	ABL81903	Human ovarian can
9	36.4	6.3	330	AAV7224	Human anti-GPIIb/I

c 10	36.4	6.3	788	22	AAH07782	Human cDNA clone (
c 11	36.4	6.3	2497	22	AAI93910	Human stomach can
c 12	36.4	6.3	2497	22	AAH18071	Human cDNA sequenc
c 13	36.4	6.3	3290	21	AAC59813	Human secreted pro
c 14	36.4	6.3	3298	22	AAS40920	cDNA encoding nove
c 15	36.2	6.3	450	23	ABV38197	Human prostate exp
c 16	35.6	6.3	3565	22	AAH40928	cDNA encoding nove
c 17	35.6	6.2	4784	22	AAH98417	Murine EST-derived
c 18	35.6	6.2	4791	22	AAS22677	Human cDNA encodin
c 19	35.4	6.2	10336	22	ABA17315	Human nervous syst
c 20	35.4	6.2	20813	22	ABA17314	Human nervous syst
c 21	35.2	6.1	4428	22	AAH06574	Bovine alpha(III)
c 22	35.2	6.1	4428	22	AAH06575	Bovine alpha(III)
c 23	35	6.1	1576	21	AAC77136	Human ORF ORF2691
c 24	34.6	6.0	1521	23	AAH71133	DNA encoding novel
c 25	34.6	6.0	2005	24	ABL95574	Human angiogenesis
c 26	34.6	6.0	2005	24	ABL88085	Human PRO7223 cDNA
c 27	34.6	6.0	2024	21	AAA07697	Human collectin en
c 28	34.6	6.0	2181	22	AAI60628	Human polynucleoti
c 29	34.6	6.0	2262	24	ABA97932	Human scavenger re
c 30	34.6	6.0	2318	22	AAI58842	Human polynucleoti
c 31	34.6	6.0	2628	22	AAH43036	Nucleotide sequenc
c 32	34.6	6.0	2641	22	AAC68903	Human EXMAD-14 cod
c 33	34.6	6.0	2929	20	AAQ92072	Human polynucleoti
c 34	34.6	6.0	2930	20	AAV55746	Human secreted pro
c 35	34.6	6.0	3535	22	AAH47802	Chimeric CBD-fused
c 36	34.6	6.0	5692	22	AAF32248	Streptomyces sp. c
c 37	34.4	6.0	2768	24	ABN89352	Mouse E-cadherin e
c 38	34.4	6.0	2768	24	ABN89352	Mouse E-cadherin e
c 39	34.2	5.9	268	13	AAQ25135	DRI. Synthetic.
c 40	34.2	5.9	874	14	AAC62307	Contig HEPATOX7, a
c 41	34	5.9	3474	14	AAQ39018	Sequence of pure m
c 42	34	5.9	3474	18	AAT74012	Maize optimised-B.
c 43	34	5.9	3474	19	AAV16191	Maize optimised DN
c 44	34	5.9	11244	22	AAK74606	Human immune/haema
c 45	34	5.9	15296	22	AAK74605	Human immune/haema

ALIGNMENTS

RESULT 1
ABQ55003
ID ABQ55003 standard; cDNA; 1594 BP.
XX AC ABQ55003;
XX AC ABQ55003;
XX DT 22-AUG-2002 (first entry)
XX DE Human ovarian antigen HSABJ44 cDNA, SEQ ID NO:883.
XX DE Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW Ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW Infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW Inflammatory condition; immune disorder; blood disorder;
KW Cardiovascular disorder; respiratory disorder; neurological disorder;
KW Gastrointestinal disorder; urinary system disorder; drug screening;
KW Gene therapy; chromosome mapping; forensic analysis;
KW Antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW Antinflammatory; gynaecological; reproductive; chromosome 12p13;
KW gene; ss.
XX OS Homo sapiens.
XX PN WO200200677-A1.
XX PD 03-JAN-2002.
XX PF 07-JUN-2001; 2001WO-US18569.
XX PR 07-JUN-2000; 2000US-209467P.

(HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;
DR WPI: 2002-147878/19.
DR P-PSDB; ABP41926.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -

PS Claim 1; SEQ ID No 883; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1594 BP; 391 A; 488 C; 434 G; 280 T; 1 other;

Query Match 85.9%; Score 494; DB 24; Length 1594;
Best Local Similarity 100.0%; Pred. No. 1.5e-127;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 GGATCGCTGCTCAAGAGCGGTCCGACGGGAGAGGACCCCAATCTGTAGCTGGAAGCTGG 141
DB 395 GGATCGCTGCTCAAGAGCGGTCCGACGGGAGAGGACCCCAATCTGTAGCTGGAAGCTGG 454
QY 142 GAGCCTCCGAGGCCATCCATCTCTCCCTGACTGTGTACAGCCACTGTACCCATTCT 201
DB 455 GAGCCTCCGAGGCCATCCATCTCTCCCTGACTGTGTACAGCCACTGTACCCATTCT 514
QY 202 GGATGTTTCCCGAGTATCCATCTGCGCTCCCGCAGCCCGAGTTTGGAGGAGGGGTG 261
DB 515 GGATGTTTCCCGAGTATCCATCTGCGCTCCCGCAGCCCGAGTTTGGAGGAGGGGTG 574
QY 262 CCGCAACAGCAGAGTCTCTGGACCTCACCAGGAGCGGACGAGTTGGAAACCGGGGACGAG 321
DB 575 CCGCAACAGCAGAGTCTCTGGACCTCACCAGGAGCGGACGAGTTGGAAACCGGGGACGAG 634
QY 322 AGCCAGGTGCCCGCAGGTACCAATGGGATTCATGTACCCGCGGGGTCTATGACTATCACT 381
DB 635 AGCCAGGTGCCCGCAGGTACCAATGGGATTCATGTACCCGCGGGGTCTATGACTATCACT 694
QY 382 GGCACATCTACATCTACATGACACAGTACTGGGGGACACCGGGTCTCTGGAGACCTC 441
XX

DB 695 GGCAACATCTACATCTACATGGACCACTACTGGGGGGACCAACCGGTCTCTGGAGACCTC 754
QY 442 CCAGCTACCCCGGAACTCCATACCCATTCGGAAGAGGGGACCTGCGCCCTCCCGGG 501
DB 755 CCAGCTACCCCGGAACTCCATACCCATTCGGAAGAGGGGACCTGCGCCCTCCCGGG 814
QY 502 CTCTCTACACCCACCCAGGAGGAGGCAAGCTTTGGCAAGCTTAGCGGAGACAGACACTGT 561
DB 815 CTCTCTACACCCACCCAGGAGGAGGCAAGCTTTGGCAAGCTTAGCGGAGACAGACACTGT 874
QY 562 GTTGCCACACCCCTC 575
DB 875 GTTGCCACACCCCTC 888

RESULT 2

ABN95637

ID ABN95637 standard; DNA; 2136 BP.

XX AC ABN95637;

XX DT 13-AUG-2002 (first entry)

XX DE Gene #2135 used to diagnose liver cancer.

XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumor; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX OS Homo sapiens.

XX PN WO200229103-A2.

XX PD 11-APR-2002.

XX PF 02-OCT-2001; 2001WO-US30589.

XX PR 02-OCT-2000; 2000US-237054P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX DR WPI; 2002-426119/45.

XX PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient, in a
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -

XX PS Claim 1; SEQ ID NO 2135; 298pp; English.

XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN95637-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 85.9%; Score 494; DB 24; Length 2136;

Best Local Similarity 100.0%; Pred. No. 1.5e-127;

Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGCCCAATCTGTAGCTGGAAGCTGG 141
|||||
Db 943 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGCCCAATCTGTAGCTGGAAGCTGG 1002
|||||
QY 142 GAGCCTCGAAGGCCCATCATCTTCCCTGACTTGTGACAGCCACTGCTACCCATTCT 201
|||||
Db 1003 GAGCCTCGAAGGCCCATCATCTTCCCTGACTTGTGACAGCCACTGCTACCCATTCT 1062
|||||
QY 202 GGAGATGTTTCCCGACATCATCTGGCTCCCGCAGCCCGGCTTTGGAGGAGGGTG 261
|||||
Db 1063 GGAGATGTTTCCCGACATCATCTGGCTCCCGCAGCCCGGCTTTGGAGGAGGGTG 1122
|||||
QY 262 CCACAACAGCAGAGTCTCTGACCTGACCAAGGAGCGCGAGTTGGAACCGGGGAGCAG 321
|||||
Db 1123 CCGCAACAGCAGAGTCTCTGACCTGACCAAGGAGCGCGAGTTGGAACCGGGGAGCAG 1182
|||||
QY 322 AGCCAGTGGCCGACGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 381
|||||
Db 1183 AGCCAGTGGCCGACGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1242
|||||
QY 382 GGCAACATCTACATCTACATGACCAAGTGGGGGAGCCCGGCTCTGGAGACCTC 441
|||||
Db 1243 GGCAACATCTACATCTACATGACCAAGTGGGGGAGCCCGGCTCTGGAGACCTC 1302
|||||
QY 442 CCAGCTACCCCGAAGCTCCATACCCCATTCGCGAAGAGGGGAGCCCTGGCCCTCCCGGG 501
|||||
Db 1303 CCAGCTACCCCGAAGCTCCATACCCCATTCGCGAAGAGGGGAGCCCTGGCCCTCCCGGG 1362
|||||
QY 502 CTCTCTACACCCGACAGGAGATGGCAAGCTTGGGACCTAGCGGAGACAGACACTGT 561
|||||
Db 1363 CTCTCTACACCCGACAGGAGATGGCAAGCTTGGGACCTAGCGGAGACAGACACTGT 1422
|||||
QY 562 GGTGCCACACCCCTC 575
|||||
Db 1423 GGTGCCACACCCCTC 1436
|||||

RESULT 3

ABK64502

ID ABK64502 standard; DNA; 2136 BP.

XX
AC ABK64502;

XX 18-JUN-2002 (first entry)

XX Human benign prostatic hyperplasia gene #397.

DE Human benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
XX Homo sapiens.

XX Wo200212440-A2.

XX 14-FEB-2002.

XX 07-AUG-2001; 2001WO-US24708.

XX 07-AUG-2000; 2000US-223323P.

XX 05-JUN-2001; 2001US-0873319.

XX (GENE-) GENE LOGIC INC.

XX (NISB) JAPAN TOBACCO INC.

XX Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;

XX WPI; 2002-257476/30.

XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by

XX detecting expression levels of one or more genes in prostate cells from

XX patient that are differentially regulated compared to normal prostate

XX cells

XX

PS Disclosure; Page 229-230; 444pp; English.

XX The invention relates to a method of diagnosing (I) the onset or
CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
CC identifying an agent that modulates the onset or progression of BPH.
CC The method is based on changes in gene expression in BPH tissue isolated
CC from patients exhibiting different clinical states of prostate
CC hyperplasia as compared to normal prostate tissue. (I) comprises
CC detecting the expression levels of one or more genes in prostate cells
CC from the subject that are differentially regulated compared to normal
CC prostate cells. (II) comprises preparing a first gene expression profile
CC of BPH cells or BPH-like cell population, exposing the cells to the
CC agent, preparing a second gene expression profile of the agent exposed
CC cells, and comparing the first and second gene expression profiles.
CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
CC useful for identifying an agent that modulates the onset or progression
CC of BPH. The methods are useful to present information identifying
CC the expression level in a tissue or cells, by comparing the expression
CC level of genes given in the specification in the tissue or cells to the
CC level of expression of gene in the database, and displaying the
CC expression levels of at least one gene in the tissue or cell sample
CC compared to the expression level in BPH. Agents using (II) are useful for
CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
CC benign prostatic hyperplasia gene sequences of the invention.

XX
SQ Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

Query Match 85.9%; Score 494; DB 24; Length 2136;

Best Local Similarity 100.0%; Pred. No. 1.6e-127;

Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGCCCAATCTGTAGCTGGAAGCTGG 141

Db 943 GGATCGCTGCTCAAGAGGCGTCCGAGGAGAGGAGGCCCAATCTGTAGCTGGAAGCTGG 1002

QY 142 GAGCCTCGAAGGCCCATCATCTTCCCTGACTTGTGACAGCCACTGCTACCCATTCT 201

Db 1003 GAGCCTCGAAGGCCCATCATCTTCCCTGACTTGTGACAGCCACTGCTACCCATTCT 1062

QY 202 GGAGATGTTTCCCGACATCATCTGGCTCCCGCAGCCCGGCTTTGGAGGAGGGTG 261

Db 1063 GGAGATGTTTCCCGACATCATCTGGCTCCCGCAGCCCGGCTTTGGAGGAGGGTG 1122

QY 262 CCACAACAGCAGAGTCTCTGACCTGACCAAGGAGCGCGAGTTGGAACCGGGGAGCAG 321

Db 1123 CCGCAACAGCAGAGTCTCTGACCTGACCAAGGAGCGCGAGTTGGAACCGGGGAGCAG 1182

QY 322 AGCCAGTGGCCGACGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 381

Db 1183 AGCCAGTGGCCGACGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1242

QY 382 GGCAACATCTACATCTACATGACCAAGTGGGGGAGCCCGGCTCTGGAGACCTC 441

Db 1243 GGCAACATCTACATCTACATGACCAAGTGGGGGAGCCCGGCTCTGGAGACCTC 1302

QY 442 CCAGCTACCCCGAAGCTCCATACCCCATTCGCGAAGAGGGGAGCCCTGGCCCTCCCGGG 501

Db 1303 CCAGCTACCCCGAAGCTCCATACCCCATTCGCGAAGAGGGGAGCCCTGGCCCTCCCGGG 1362

QY 502 CTCTCTACACCCGACAGGAGATGGCAAGCTTGGGACCTAGCGGAGACAGACACTGT 561

Db 1363 CTCTCTACACCCGACAGGAGATGGCAAGCTTGGGACCTAGCGGAGACAGACACTGT 1422

QY 562 GGTGCCACACCCCTC 575

Db 1423 GGTGCCACACCCCTC 1436

RESULT 4

ABL64078

ID ABL64078 standard; DNA; 2136 BP.

XX

AC ABL64078;

XX	15-MAY-2002	(first entry)	
DT	Breast cancer related gene sequence	SEQ ID NO:2415.	
XX	Human; cancer; colon; breast; ovary; cesophagus; kidney; thyroid;		
XX	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;		
KW	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;		
KW	gene; ds.		
XX	Homo sapiens.		
OS	WO2001194629-A2.		
PN	13-DEC-2001.		
XX	30-MAY-2001; 2001WO-US10838.		
XX	05-JUN-2000; 2000US-209473P.		
XX	05-JUN-2000; 2000US-209531P.		
PR	18-SEP-2000; 2000US-233133P.		
PR	18-SEP-2000; 2000US-233617P.		
PR	20-SEP-2000; 2000US-234009P.		
PR	20-SEP-2000; 2000US-234034P.		
PR	20-SEP-2000; 2000US-234052P.		
PR	22-SEP-2000; 2000US-234509P.		
PR	22-SEP-2000; 2000US-234567P.		
PR	25-SEP-2000; 2000US-234923P.		
PR	25-SEP-2000; 2000US-234924P.		
PR	25-SEP-2000; 2000US-235077P.		
PR	25-SEP-2000; 2000US-235082P.		
PR	25-SEP-2000; 2000US-235134P.		
PR	25-SEP-2000; 2000US-235280P.		
PR	26-SEP-2000; 2000US-235637P.		
PR	26-SEP-2000; 2000US-235638P.		
PR	27-SEP-2000; 2000US-235711P.		
PR	27-SEP-2000; 2000US-235720P.		
PR	27-SEP-2000; 2000US-235840P.		
PR	27-SEP-2000; 2000US-235863P.		
PR	28-SEP-2000; 2000US-236028P.		
PR	28-SEP-2000; 2000US-236032P.		
PR	28-SEP-2000; 2000US-236033P.		
PR	28-SEP-2000; 2000US-236034P.		
PR	28-SEP-2000; 2000US-236109P.		
PR	28-SEP-2000; 2000US-236111P.		
PR	29-SEP-2000; 2000US-236842P.		
PR	29-SEP-2000; 2000US-236891P.		
PR	02-OCT-2000; 2000US-237172P.		
PR	02-OCT-2000; 2000US-237173P.		
PR	02-OCT-2000; 2000US-237278P.		
PR	02-OCT-2000; 2000US-237294P.		
PR	02-OCT-2000; 2000US-237295P.		
PR	02-OCT-2000; 2000US-237316P.		
PR	03-OCT-2000; 2000US-237425P.		
PR	03-OCT-2000; 2000US-237598P.		
PR	03-OCT-2000; 2000US-237604P.		
PR	03-OCT-2000; 2000US-237606P.		
PR	03-OCT-2000; 2000US-237608P.		
PR	01-NOV-2000; 2000US-244867P.		
PR	01-NOV-2000; 2000US-245084P.		
XX	(AVAL-) AVALON PHARM.		
XX	Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;		
PI	Soppet DR, Weaver Z;		
PI	WPI; 2002-188264/24.		
XX	Screening for anti-neoplastic agent involves exposing cells to a		
XX	chemical agent to be tested for anti-neoplastic activity, and		
PT	determining a change in expression of a gene of a signature gene set		
PT	Claim 1; SEQ ID 2415; 44pp; English.		
XX			

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, esophagael, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 2136 BP; 446 A; 706 C; 608 G; 376 T; 0 other;

QY 443 CAGCTACCCCGGACCTCCATACCCCAT 470
|||||
DB 49 CAGCTACCCCGGACCTCCATTCGCCCT 22

RESULT 8

ABL81903
ID ABL81903 standard; cDNA; 481 BP.

XX AC ABL81903;
XX 17-MAY-2002 (first entry)
XX DE Human ovarian cancer related cDNA clone SEQ ID NO:4881.

XX DE Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX KW Homo sapiens.
XX OS WO200192581-A2.

XX PN 06-DEC-2001.

XX PD 29-MAY-2001; 2001WO-US17756.

XX PF 26-MAY-2000; 2000US-207484P.

XX PR (CORI-) CORIXA CORP.

XX PA Algate PA, Harlocker SL, Jones R;

XX PI WPI; 2002-122075/16.

XX DR Composition for therapy and diagnosis of ovarian cancer comprising
XX PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
XX PT polypeptide, antibody specific to polypeptide or T cell expressing
XX PT polypeptide -
XX PS Claim 1; SEQ ID 4881; 489pp; English.

XX CC The present invention describes a composition (I) comprising: carriers
XX CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
XX CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
XX CC (SI) from the 10912 nucleotide sequences as given in ABL77023 to
XX CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
XX CC population of (II), or antigen presenting cells that express (II).
XX CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
XX CC (SI) can be used for detecting ovarian cancer in a patient's biological
XX CC sample preferably serum or ovarian tissue. The method comprises
XX CC contacting a biological sample from a patient with (IV), detecting the
XX CC amount of polynucleotide hybridising to (IV) and comparing the amount to
XX CC a predetermined cutoff value and thereby detecting ovarian cancer in the
XX CC patient, where the amount of polynucleotide hybridising to (IV) is
XX CC detected preferably by polymerase chain reaction (PCR). (I) comprising
XX CC (III) and/or (II) is useful for stimulating and/or expanding T cells
XX CC specific for an ovarian tumour protein comprising contacting T cells
XX CC with (III) or (II). (III) is useful in design and preparation of
XX CC ribozyme molecules for inhibiting expression of the tumour polypeptides
XX CC and proteins in tumour cells; and to isolate a full length gene from a
XX CC suitable library e.g., a tumour cDNA library using well known
XX CC techniques.

SQ Sequence 481 BP; 113 A; 161 C; 123 G; 84 T; 0 other;

Query Match 10.6%; Score 61.2; DB 24; Length 481;
Best Local Similarity 95.5%; Pred. NO. 2.9e-07;
Matches 63; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CTGGTGGAGGAGCTCCAGGCACTGCCGACACCACTGCAGAAATCCATTAGAG 60
|||||

DB 416 CTGGTGGAGGAGATCCAGGCACTGCACAGTCCACACCACTGCAGAAATCCATTAGAG 475

QY 61 CCACTG 66

Db 476 CCACTG 481
|||||

RESULT 9

AAV72224
ID AAV72224 standard; DNA; 333 BP.

XX AC AAV72224;
XX 07-SEP-1999 (first entry)
XX DE Human anti-GPIIb/IIIa antibody light chain DNA from phagemid PDG7.

XX DE Antibody; GPIIb/IIIa; human; auto-antibody; anti-idiotypic; diagnosis;
XX KW blood platelet membrane protein; predisposition; prevention; treatment;
XX KW autoimmune thrombocytopaenic purpura; AITP; fibrinogen binding; thrombi;
XX KW thrombocyte; cardiac infarction; pulmonary embolism; light chain; ds.

XX OS Homo sapiens.

XX PH Key
XX PH misc_feature
XX FT Location/Qualifiers
XX FT 1..60
XX FT /tag= a
XX FT /note= "Framework region 1 (FR1)"

XX FT 61..99
XX FT /tag= b
XX FT /note= "complementarity determining region 1 (CDR1)"

XX FT 100..144
XX FT /tag= c
XX FT /note= "Framework region 2 (FR2)"

XX FT 145..165
XX FT /tag= d
XX FT /note= "complementarity determining region 2 (CDR2)"

XX FT 166..261
XX FT /tag= e
XX FT /note= "Framework region 3 (FR3)"

XX FT 262..294
XX FT /tag= f
XX FT /note= "complementarity determining region 3 (CDR3)"

XX FT 295..333
XX FT /tag= g
XX FT /note= "Framework region 4 (FR4)"

XX FT WO9855619-A1.
XX PD 10-DEC-1998.

XX PF 05-JUN-1998; 98WO-EP03397.

XX PR 08-MAY-1998; 98DE-1020663.

XX PR 06-JUN-1997; 97DE-1023904.

XX PR 12-DEC-1997; 97DE-1055227.

XX PA (ASAT-) ASAT AG APPLIED SCI & TECHNOLOGY.
XX PI Berchtold P, Escher RPA;

XX DR WPI; 1999-105496/09.

XX DR P-PSDB; AAW90279.

XX PT Nucleic acid encoding human autoantibodies against platelet
XX PT glycoprotein IIb/IIIa - used for diagnosis, treatment and prevention
XX PT of autoimmune thrombocytopaenic purpura and for modulation of
XX PT fibrinogen binding
XX PS Disclosure; Page 49-50; 93pp; German.

XX CC This invention describes novel nucleic acid fragments that encode human
XX CC auto-antibodies and anti-idiotypic antibodies against blood platelet
XX CC membrane protein, GPIIb/IIIa. The products of the invention are used
XX CC for diagnosis (including monitoring and determining predisposition),
XX CC prevention and treatment of autoimmune thrombocytopaenic purpura (AITP)

CC and also for modulating binding of fibrinogen to thrombocytes
 CC (particularly to dissolve thrombi and/or prevent their formation, e.g.
 CC in cases of cardiac infarction or pulmonary embolism). Unlike murine
 CC antibodies, human antibodies (hAb) do not induce adverse side effects
 CC and persist for longer in vivo than small peptides.
 XX
 SQ Sequence 333 BP; 59 A; 109 C; 93 G; 72 T; 0 other;
 Query Match 6.3%; Score 36.4; DB 20; Length 333;
 Best Local Similarity 49.0%; Pred. No. 2.1;
 Matches 97; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
 QY 5 TGGAGGAGCTCCAGGAGCTCCGACGACCAACCTGCAAAATCCATTAGAGCCAC 64
 Db TGGAGGAGCTCCAGGAGCTCCGACGACCAACCTGCTAGCTGATCACCAGGTCCAGGCAC 122
 QY 65 TGCCCCCAGAGATGTCAGGATCGTCTCAAGAGCGCTCCGAGGGAGAGGGACCCCAATC 124
 Db TGGAGGAGCTCCAGGAGCTCCGACGACCAACCTGCTAGCTGATCACCAGGTCCAGGCAC 182
 QY 123 GGCCCCCAACTCTCATCTTTGGTAGTCATCAGCGGCCCTCAGGGGTCCCTGACCGATT 182
 QY 125 CTCTAGCTGGAAGCTGGGAGCCTCCGAAGGCCCATCTTCCCTGACTTGGTACAGC 184
 Db CTCTGGCTCAAGTCGGGACCTCCGCTCCCTGGCCATCGTGGGTCCCAATCTGGGA 242
 QY 185 CACTGCTACCCATTCTG 202
 Db TGCTGGTGACTATTACTG 260
 RESULT 10
 AAH07782/c
 ID AAH07782 standard; cDNA; 788 BP.
 XX
 AC AAH07782;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:4617.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 1; SEQ ID 4617; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 788 BP; 179 A; 204 C; 249 G; 153 T; 3 other;
 Query Match 6.3%; Score 36.4; DB 22; Length 788;
 Best Local Similarity 51.2%; Pred. No. 2.7;
 Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
 QY 404 GACAGTACTTGGGGGACCAACCGGTCCTGGAGACCTCCAGCTACCCCGAACCTCCAT 463
 Db TGGAGGAGCTCCAGGAGCTCCGACGACCAACCTGCTAGCTGATCACCAGGTCCAGGCAC 355
 QY 464 ACCCATATCCGGAAGAGGGGACCTCCGCGCTCCCGGGTCTCTACACCCACCAAGAG 523
 Db TTTCTATCATGAGAGCGGGGACGATGCCAAAGAGGGCTCATCGCCATGGCAGGGAGG 295
 QY 524 ATGGCAAGGCTTGGCACCTAGCGAGACAGACGACACTGTGGTGGCCAC 569
 Db ATTTCCGCCCTTCTTCCGAGGGCGTGGTGGCATGAGATGCCAC 249
 RESULT 11
 AA193910/c
 ID AA193910 standard; cDNA; 2497 BP.
 XX
 AC AA193910;
 XX
 DT 13-NOV-2001 (first entry)
 XX
 DE Human stomach cancer expressed polynucleotide SEQ ID NO 134.
 XX
 KW Human; stomach cancer; marker; screening; micro-metastasis;
 KW peritoneal dissemination; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200109317-A1.
 XX
 PD 08-FEB-2001.
 XX
 PF 28-JUL-2000; 2000WO-JP05063.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 18-OCT-1999; 99US-0159590.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 17-FEB-2000; 2000US-0183322.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Aburatani H;
 PI Kodama T, Midorikawa Y;
 XX

DR WPI; 2001-570287/64.
XX P-PSDB; AAM94032.
XX New Stomach cancer-associated genes, useful as markers in blood tests
PT for screening for the early stages of the disease -
XX
XX
XX Claim 1; Page 216-217; 242pp; Japanese.
XX
XX The invention relates to stomach cancer-expressed genes
CC (AAI93842-AAI93917) and the encoded proteins (AAM93967-AAM94039). The
CC genes can be used as markers in blood tests for screening for the early
CC stages of the disease. The proteins and peptides can be used as targets
CC for screening for compounds to treat the disease. They can also be used
CC for predicting micro-metastases. The gene can predict peritoneal
CC dissemination.
XX
XX Sequence 2497 BP; 543 A; 605 C; 735 G; 614 T; 0 other;
SQ
Query Match 6.3%; Score 36.4; DB 22; Length 2497;
Best Local Similarity 51.2%; Pred. No. 4;
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 404 GACCAGTACTGGGGGACACCGGGTCTGGAGACCTCCAGCTACCCCGGACCTCCAT 463
DB 414 GGGCAGGCTGGGAGATGCACAGGGCCCGGAGACGTTGCTGCTCCACGCGTGCCC 355
QY 464 ACCCATTCGCCGAAGAGGGGGACCTGGCCCTCCCGGGCTCTACACCCACAGGAAG 523
DB 354 TTCTCATCTCAGGACGGGGAGGATGCCAAGAGGGGCTCATCGCCATGGCAGGGAGG 295
QY 524 ATGGCAAGGCTTGGCACCTTAGCGGAGACAGACACTGTGGTGCCAC 569
DB 294 ATTTCCGCCCTTCTTCCGAGGGCCGTGTGCGATGCAGATGCCAC 249
RESULT 12
AAH18071/c
ID AAH18071 standard; cDNA; 2497 BP.
XX
XX AAH18071;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA sequence SEQ ID NO:17912.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX

PS Claim 8; SEQ ID 17912; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
SQ Sequence 2497 BP; 543 A; 605 C; 735 G; 614 T; 0 other;
Query Match 6.3%; Score 36.4; DB 22; Length 2497;
Best Local Similarity 51.2%; Pred. No. 4;
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 404 GACCAGTACTGGGGGACACCGGGTCTGGAGACCTCCAGCTACCCCGGACCTCCAT 463
DB 414 GGGCAGGCTGGGAGATGCACAGGGCCCGGAGACGTTGCTGCTCCACGCGTGCCC 355
QY 464 ACCCATTCGCCGAAGAGGGGGACCTGGCCCTCCCGGGCTCTACACCCACAGGAAG 523
DB 354 TTCTCATCTCAGGACGGGGAGGATGCCAAGAGGGGCTCATCGCCATGGCAGGGAGG 295
QY 524 ATGGCAAGGCTTGGCACCTTAGCGGAGACAGACACTGTGGTGCCAC 569
DB 294 ATTTCCGCCCTTCTTCCGAGGGCCGTGTGCGATGCAGATGCCAC 249
RESULT 13
AAC59813/c
ID AAC59813 standard; DNA; 3290 BP.
XX
XX AAC59813;
XX
XX 26-JAN-2001 (first entry)
XX
XX Human secreted protein encoding DNA clone vo9 1.
XX
XX Secreted protein; human; autoimmune disorder; multiple sclerosis; ulcer;
KW systemic lupus erythematosus; rheumatoid arthritis; anaemia; stroke;
KW haematopoiesis regulation; tissue regrowth; wound healing; haemophilia;
KW Alzheimer's disease; Parkinson's disease; Shy-drager syndrome; cancer;
KW contraceptive; infection; growth inhibition; hyperproliferative disorder;
KW psoriasis; ds.
XX
XX Homo sapiens.
XX
XX WO200055375-A1.
XX
XX 21-SEP-2000.
XX
XX 17-MAR-2000; 2000WO-US07285.
XX
XX 17-MAR-1999; 99US-0124808.
XX
XX 17-MAR-1999; 99US-0124916.
XX
XX 17-AUG-1999; 99US-0149639.
XX
XX 01-OCT-1999; 99US-0157247.
XX

DT 16-SEP-2002 (first entry)
 XX Human prostate expression marker cdna 38188.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.
 XX
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer -
 XX
 PS Claim 1; Page 7795; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 450 BP; 90 A; 130 C; 127 G; 103 T; 0 other;
 Query Match 6.3%; Score 36.2; DB 23; Length 450;
 Best Local Similarity 51.8%; Pred. No. 2.6;
 Matches 83; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
 QY 79 TCAGGATCGTGTCTCAAGAGCGCTCCGAGGAGAGGAGGCCCAATCCTGTAGCTGGAAGC 138
 DB 26 TCTAGATGATGCTCGAGGCGGCCGACGTGTGATGATATCTGCAGAAATTCGCCCTTAGC 85
 QY 139 TGGAGGCTCGAAGGCCCATTCATCTCCCTGACTGTGTGATAGCCACTGCTACCCATT 198
 DB 86 TGTGCGGCGGCGAGGTACCTTCTCTCCCTCCCTGCTGTAACATCATTCAGAACAGGAA 145
 QY 199 TCTGGAGATGTTTCCCCAGTATCCACTGGGCTCCCGCCAGC 239
 DB 146 CATGGGAAGCAGCCAGCCCTGCTGAGCCGGGAGGAGC 186

Search completed: April 15, 2003, 15:16:36
 Job time : 112.083 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:09:45 ; Search time 664.156 Seconds
(without alignments)
14021.414 Million cell updates/sec

Title: US-09-917-372-2_COPY_804_1378

Perfect score: 575

Sequence: 1 ctggtggaggcagctccagg.....cactgtggtgccacaccctc 575

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estcov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	85.9	814	9	AU118203
2	494	85.9	868	10	BE275240
3	494	85.9	895	14	BQ957322
4	494	85.9	920	12	BG036632
5	494	85.9	954	13	BI821009
6	494	85.9	968	14	BQ898015

7	482	83.8	714	13	BI829057
8	482	83.8	901	12	BE740143
9	481.4	83.7	954	14	BQ930450
10	480.4	83.5	900	13	BI769498
11	452.8	78.7	749	12	BE789430
12	445	77.4	735	12	BG331666
13	444.8	77.4	606	13	BI905922
14	437.2	76.0	700	12	BG576875
15	432	75.1	598	12	BG403820
16	425.4	74.0	427	10	AW630662
17	423	73.6	801	12	BF792867
18	417	72.5	458	14	BM766807
19	417	72.5	537	14	BM766352
20	417	72.5	572	14	BM767287
21	417	72.5	666	14	BM767318
22	416	72.3	501	14	BM766833
23	415.4	72.2	634	14	BM746911
24	412	71.7	598	14	BM723050
25	409.4	71.2	579	10	BE207591
26	381	66.3	640	12	BG491235
27	376	65.4	506	14	BM855039
28	375	65.2	808	13	BI818411
29	374.4	65.1	466	14	BQ082995
30	370.8	64.5	426	14	BM856797
31	362.8	63.1	416	14	BQ083004
32	361.2	62.8	418	14	BM855410
33	350.2	60.9	942	13	BI818552
34	347.6	60.5	721	9	AI336202
35	344.8	60.0	360	14	BM766382
36	337	58.6	706	12	BG698325
37	332.2	57.8	633	9	AI870838
38	326	56.7	728	13	BI765703
39	317.2	55.2	530	12	BG610053
40	315.4	54.9	386	9	AA099514
41	314	54.6	882	14	BQ691803
42	306.8	53.4	518	13	BI338842
43	282	49.0	598	14	BQ305454
44	278.8	48.5	665	13	BI554834
45	278.8	48.5	934	14	BQ928227

ALIGNMENTS

RESULT 1	AU118203	814 bp	linear	EST 01-AUG-2002
LOCUS	AU118203	HEMBA1	Homo sapiens	cdna clone HEMBA1003089 5', mRNA
DEFINITION	AU118203	sequence.		
ACCESSION	AU118203			
VERSION	AU118203.1	GI:10933231		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.			
TITLE	HRI human cdna project			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3975 Fax: 81-438-52-3986 Email: genomics@hri.co.jp HRI human cdna project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cdna library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.			

```

FEATURES
  source
    Location/Qualifiers
      1..814
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="HEMBA1003089"
        /clone_lib="HEMBA1"
        /tissue_type="whole embryo, mainly head"
        /dev_stage="embryo, 10 weeks"
        /note="Vector: pME18SFL3"
      173 a 255 c 232 g 151 t
      3 others
BASE COUNT 173 a 255 c 232 g 151 t
ORIGIN
  Query Match 85.9%; Score 494; DB 9; Length 814;
  Best Local Similarity 100.0%; Pred. No. 1.3e-111;
  Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 GGATCGCTCTCAAGAGGGCTCCGAGGAGGAGGACCAATCCCTAGCTGGAAGCTGG 141
Db 98 GGATCGCTCTCAAGAGGGCTCCGAGGAGGAGGACCAATCCCTAGCTGGAAGCTGG 157
QY 142 GAGCCTCCGAAGGCCCATCATCACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 201
Db 158 GAGCCTCCGAAGGCCCATCATCACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 217
QY 202 GGAGATGTTTCCCACTATCACTGGCTCCCGGAGGCCCAAGTTTGGAGGAGGGGTG 261
Db 218 GGAGATGTTTCCCACTATCACTGGCTCCCGGAGGCCCAAGTTTGGAGGAGGGGTG 277
QY 262 CGCAACAGCAGAGTCTCTGACCTGACCGAGGAGGAGGAGGAGGAGGAGGAGGAG 321
Db 278 CGCAACAGCAGAGTCTCTGACCTGACCGAGGAGGAGGAGGAGGAGGAGGAGGAG 337
QY 322 AGCCAGGTGGCCCGTACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 381
Db 338 AGCCAGGTGGCCCGTACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 397
QY 382 GGCACATCTACATCTACATGACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 441
Db 398 GGCACATCTACATCTACATGACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 457
QY 442 CCAGTACCCCGAAGTCCATACCCCATTCGCAAGAGGGGAGCCCTGGCCCTCCGGG 501
Db 458 CCAGTACCCCGAAGTCCATACCCCATTCGCAAGAGGGGAGCCCTGGCCCTCCGGG 517
QY 502 CTCTCTACACCCCGAAGATGGCAAGGCTTGGCAAGTGGCAAGTGGCAAGTGGCAAGT 561
Db 518 CTCTCTACACCCCGAAGATGGCAAGGCTTGGCAAGTGGCAAGTGGCAAGTGGCAAGT 577
QY 562 GGTGCCACACCCCTC 575
Db 578 GGTGCCACACCCCTC 591
RESULT 2
BE275240 868 bp mRNA linear EST 13-JUL-2000
LOCUS 601122062P1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3346311 5',
DEFINITION mRNA sequence.
ACCESSION BE275240
VERSION BE275240.1 GI:9150193
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 868)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory

```

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LICM135 row: k column: 16
High quality sequence stop: 746.
FEATURES
  Location/Qualifiers
    1..868
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="IMAGE:3346311"
      /clone_lib="NIH_MGC_20"
      /tissue_type="melanotic melanoma"
      /lab_host="DH10B (phage-resistant)"
      /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2:
      EcoRI; cDNA made by oligo-dT priming. Directionally
      cloned into EcoRI/XhoI sites using the following 5'
      adaptor: GGCACGAG(G). Size-selected >500bp for average
      insert size 1.8kb. Library constructed by Ling Hong in
      the laboratory of Gerald M. Rubin (University of
      California, Berkeley) using ZAP-cDNA synthesis kit
      (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 170 a 286 c 253 g 159 t
ORIGIN
  Query Match 85.9%; Score 494; DB 10; Length 868;
  Best Local Similarity 100.0%; Pred. No. 1.3e-111;
  Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 GGATCGCTCTCAAGAGGGCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 141
Db 138 GGATCGCTCTCAAGAGGGCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 197
QY 142 GAGCCTCCGAAGGCCCATCATCACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 201
Db 198 GAGCCTCCGAAGGCCCATCATCACTTCCCTGACTTGGTACAGCCACTGCTACCCATTCT 257
QY 202 GGAGATGTTTCCCACTATCACTGGCTCCCGGAGGCCCAAGTTTGGAGGAGGGGTG 261
Db 258 GGAGATGTTTCCCACTATCACTGGCTCCCGGAGGCCCAAGTTTGGAGGAGGGGTG 317
QY 262 CGCAACAGCAGAGTCTCTGACCTGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 321
Db 318 CGCAACAGCAGAGTCTCTGACCTGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 377
QY 322 AGCCAGGTGGCCCGTACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 381
Db 378 AGCCAGGTGGCCCGTACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 437
QY 382 GGCACATCTACATCTACATGACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 441
Db 438 GGCACATCTACATCTACATGACCAATGGCATTCATGTCACCGGGGTCTATGACTATCACT 497
QY 442 CCAGTACCCCGAAGTCCATACCCCATTCGCAAGAGGGGAGCCCTGGCCCTCCGGG 501
Db 498 CCAGTACCCCGAAGTCCATACCCCATTCGCAAGAGGGGAGCCCTGGCCCTCCGGG 557
QY 502 CTCTCTACACCCCGAAGATGGCAAGGCTTGGCAAGTGGCAAGTGGCAAGTGGCAAGT 561
Db 558 CTCTCTACACCCCGAAGATGGCAAGGCTTGGCAAGTGGCAAGTGGCAAGTGGCAAGT 617
QY 562 GGTGCCACACCCCTC 575
Db 618 GGTGCCACACCCCTC 631

```

```

RESULT 3
BQ957322
LOCUS
DEFINITION BQ957322
ACCESSION BQ957322
VERSION BQ957322.1 GI:22372800

```

```

BQ957322 895 bp mRNA linear EST 21-AUG-2002
AGENCOURT-8777754 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6384424
5', mRNA sequence.
ACCESSION BQ957322
VERSION BQ957322.1 GI:22372800

```

Db	552	CTCTCTACACCCACACAGGAAGATGGCAAGCTTTGGCACTTAGCGGAGACAGACACTGT	611
QY	562	GGTGCCACACCTC	575
Db	612	GGTGCCACACCTC	625
RESULT 4			
LOCUS	BG036632		
DEFINITION	602326834F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4428225 5',	920 bp	mRNA linear EST 24-JAN-2001
ACCESSION	BG036632		mRNA sequence.
VERSION	BG036632.1	GI:12432013	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 920)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1AM10178 row: c column: 10 High quality sequence stop: 695. Location/Qualifiers 1..920 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4428225" /clone_lib="NIH_MGC_91" /tissue_type="adenocarcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-qt primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."		
BASE COUNT	187 a	304 c	270 g 159 t
ORIGIN			
Query Match	85.9%;	Score 494;	DB 12; Length 920;
Best Local Similarity	100.0%;	Pred. No. 1.3e-111;	
Matches 494;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	82	GGATCGTGCTCAAGAGCGCTGCCGAGGAGAGGCCCAATCTCTAGCTGGAAGCTGG	141
Db	52	GGATCGTGCTCAAGAGCGCTGCCGAGGAGAGGCCCAATCTCTAGCTGGAAGCTGG	111
QY	142	GAGCCTCCGAGGCCCATCATACTTCCCTGACTTGGTACAGCCACTGCTACCCATTTCT	201
Db	112	GAGCCTCCGAGGCCCATCCATACTTCCCTGACTTGGTACAGCCACTGCTACCCATTTCT	171
QY	202	GGAGATGTTTCCCGAGTATCCACTGGGCTCCCCCGACCCAGTTTGGAGGCGAGGGTG	261
Db	172	GGAGATGTTTCCCGAGTATCCACTGGGCTCCCCCGACCCAGTTTGGAGGCGAGGGTG	231
QY	262	CCGACAGCAGAGTCTCTGGACCTGACCGAGGAGCGCGAGTTTGGAAACCCGGGGAGCAG	321
Db	232	CCGACAGCAGAGTCTCTGGACCTGACCGAGGAGCGCGAGTTTGGAAACCCGGGGAGCAG	291
QY	322	AGCCAGGTGGCCCAACGGTACCAATGGCATTCATGTACCGCGGGGTCTATGACTATCACT	381
Db	292	AGCCAGGTGGCCCAACGGTACCAATGGCATTCATGTACCGCGGGGTCTATGACTATCACT	351

```

QY 382 GGCAACATCTACATCTACATGACAGTACTGGGGGACCAACCGGGTCTCTGGAGACCTC 441
|||||
Db 352 GGCAACATCTACATCTACATGACAGTACTGGGGGACCAACCGGGTCTCTGGAGACCTC 411
|||||
QY 442 CCAGTACCCCGAGACCTCTCATACCCATTCCTCCGAGAGGGGACCTCGGCCCTCCCGGG 501
|||||
Db 412 CCAGTACCCCGAGACCTCTCATACCCATTCCTCCGAGAGGGGACCTCGGCCCTCCCGGG 471
|||||
QY 502 CTCTCTACACCCACCAAGAGATGGCAAGCTTGGCACCTAGCCGAGACAGACAGACTGT 561
|||||
Db 472 CTCTCTACACCCACCAAGAGATGGCAAGCTTGGCACCTAGCCGAGACAGACAGACTGT 531
|||||
QY 562 GGTGCCACACCCCTC 575
|||||
Db 532 GGTGCCACACCCCTC 545
|||||

RESULT 5
BI821009
LOCUS
DEFINITION
603035664F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176935 5',
mRNA sequence.
ACCESSION
BI821009
VERSION
BI821009.1 GI:15932559
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 954)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1440 row: o column: 16
High quality sequence stop: 758.
FEATURES
Location/Qualifiers
1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5176935"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site:1; NotI; Site:2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT 212 a 313 c 281 g 148 t
ORIGIN
Query Match 85.9%; Score 494; DB 13; Length 954;
Best Local Similarity 100.0%; Pred. No. 1.3e-111;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 GGATCGCTGCTCAGAGCGCTCCGAGGAGAGGGACCAATCTGTAGCTGGAAGCTGG 141
|||||
Db 14 GGATCGCTGCTCAGAGCGCTCCGAGGAGAGGGACCAATCTGTAGCTGGAAGCTGG 73
|||||

```

```

QY 142 GAGCTCCGAAGGCCATCCATCTTCCCTGACTTGTGTACAGCCACTGCTACCCATTCT 201
|||||
Db 74 GAGCTCCGAAGGCCATCCATCTTCCCTGACTTGTGTACAGCCACTGCTACCCATTCT 133
|||||
QY 202 GGAGATGTTTCCCAAGTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGCGAGGGTG 261
|||||
Db 134 GGAGATGTTTCCCAAGTATCCACTGGGCTCCCGCAGCCCCAGTTTGGAGGCGAGGGTG 193
|||||
QY 262 CCGCAACAGCAGAGTCTCTCTGGACCTGACACGAGGAGCGGAGTGGAAACCCGGGAGCAG 321
|||||
Db 194 CCGCAACAGCAGAGTCTCTCTGGACCTGACACGAGGAGCGGAGTGGAAACCCGGGAGCAG 253
|||||
QY 322 AGCCAGGTGGCCCGGTTACCAATGGCATTCATCTACCCGGGGTCTATGACTATCACT 381
|||||
Db 254 AGCCAGGTGGCCCGGTTACCAATGGCATTCATCTACCCGGGGTCTATGACTATCACT 313
|||||
QY 382 GGCAACATCTACATCTACAATGGACAGTACTGGGGGACCAACCGGGTCTCTGGAGACCTC 441
|||||
Db 314 GGCAACATCTACATCTACAATGGACAGTACTGGGGGACCAACCGGGTCTCTGGAGACCTC 373
|||||
QY 442 CCAGTACCCCGAGACCTCTCATACCCATTCCTCCGAGAGGGGACCTCGGCCCTCCCGGG 501
|||||
Db 374 CCAGTACCCCGAGACCTCTCATACCCATTCCTCCGAGAGGGGACCTCGGCCCTCCCGGG 433
|||||
QY 502 CTCTCTACACCCACCAAGAGATGGCAAGCTTGGCACCTAGCCGAGACAGACAGACTGT 561
|||||
Db 434 CTCTCTACACCCACCAAGAGATGGCAAGCTTGGCACCTAGCCGAGACAGACAGACTGT 493
|||||
QY 562 GGTGCCACACCCCTC 575
|||||
Db 494 GGTGCCACACCCCTC 507
|||||

RESULT 6
BO898015
LOCUS
DEFINITION
BO898015 968 bp mRNA linear EST 16-AUG-2002
clone IMAGE:6177574 5', mRNA sequence.
ACCESSION
BO898015
VERSION
BO898015.1 GI:22290029
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 968)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13555 row: l column: 23
High quality sequence stop: 615.
FEATURES
Location/Qualifiers
1..968
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6177574"
/clone_lib="Lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site:1;
Note1: Site:2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:

```

5'-TCGACCCACGCGTCGG-3' and
5'-GACTAGTTCTAGATCGGAGCGCGCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies.
BASE COUNT 219 a 320 c 260 g 168 t 1 others
ORIGIN

Query Match 85.9%; Score 494; DB 14; Length 968;
Best Local Similarity 100.0%; Pred. No. 1.3e-111; Mismatches 0; Indels 0; Gaps 0;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 GGATCGCTCTCAAGAGCGTCCGAGGAGGAGGACCAATCTGTAGCTGGAAGCTGG 141
|||||
Db 266 GGATCGCTCTCAAGAGCGTCCGAGGAGGAGGACCAATCTGTAGCTGGAAGCTGG 325
|||||
QY 142 GAGCTCCGAGGCGCATCTATCTCCCTGACTTGGTACAGCCACTGTACCCATTTCT 201
|||||
Db 326 GAGCTCCGAGGCGCATCTATCTCCCTGACTTGGTACAGCCACTGTACCCATTTCT 385
|||||
QY 202 GGAGATGTTTCCCGAGTATCCACTGGCTCCCGGAGGCGGCGGCTTTGGAGCGAGGGTG 261
|||||
Db 386 GGAGATGTTTCCCGAGTATCCACTGGCTCCCGGAGGCGGCGGCTTTGGAGCGAGGGTG 445
|||||
QY 262 CCGCAACAGCAGAGTCTCTGACCTGACGAGGAGCGGAGTGGAAACCGGGGAGAG 321
|||||
Db 446 CCGCAACAGCAGAGTCTCTGACCTGACGAGGAGCGGAGTGGAAACCGGGGAGAG 505
|||||
QY 322 AGCCAGGTGGCCAGGTACCAATGGCAATCATGTACCGGGGGTCTATGACTATCACT 381
|||||
Db 506 AGCCAGGTGGCCAGGTACCAATGGCAATCATGTACCGGGGGTCTATGACTATCACT 565
|||||
QY 382 GGCACATCTACATCTACAATGGACAGTACTGGGGGAGGACCGGGTCTGGAGACCTC 441
|||||
Db 566 GGCACATCTACATCTACAATGGACAGTACTGGGGGAGGACCGGGTCTGGAGACCTC 625
|||||
QY 442 CCAGCTACCCCGAAGCTTCATACCCCATTCGCGAGAGGGGAGCCTGGCCCTCCCGGG 501
|||||
Db 626 CCAGCTACCCCGAAGCTTCATACCCCATTCGCGAGAGGGGAGCCTGGCCCTCCCGGG 685
|||||
QY 502 CTCTCTACACCCACAGAGATGGCAAGGCTTGGACCTTACGCGGAGACAGCACTGT 561
|||||
Db 686 CTCTCTACACCCACAGAGATGGCAAGGCTTGGACCTTACGCGGAGACAGCACTGT 745
|||||
QY 562 GGTGCCACACCCCTC 575
|||||
Db 746 GGTGCCACACCCCTC 759
|||||

RESULT 7
BI829057
LOCUS 603079147F1 NIH_MGC_119 Homo sapiens cdna clone IMAGE:5171122 5',
DEFINITION mRNA sequence.
ACCESSION BI829057
VERSION BI829057.1 GI:15940607
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 714)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
Library Preparation: Life Technologies, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LAM11425 row: m column: 11
High quality sequence stop: 711.
location/Qualifiers
1. 714
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5171122"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
BASE COUNT 152 a 231 c 202 g 129 t
ORIGIN

Query Match 83.8%; Score 482; DB 13; Length 714;
Best Local Similarity 99.8%; Pred. No. 1.1e-108;
Matches 493; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 82 GGATCGCTCTCAAGAGCGTCCGAGGAGGAGGACCAATCTGTAGCTGGAAGCTGG 141
|||||
Db 81 GGATCGCTCTCAAGAGCGTCCGAGGAGGAGGACCAATCTGTAGCTGGAAGCTGG 140
|||||
QY 142 GAGCTCCGAGGCGCATCTATCTCCCTGACTTGGTACAGCCACTGTACCCATTTCT 201
|||||
Db 141 GAGCTCCGAGGCGCATCTATCTCCCTGACTTGGTACAGCCACTGTACCCATTTCT 200
|||||
QY 202 GGAGATGTTTCCCGAGTATCCACTGGCTCCCGGAGGCGGCGGCTTTGGAGCGAGGGTG 261
|||||
Db 201 GGAGATGTTTCCCGAGTATCCACTGGCTCCCGGAGGCGGCGGCTTTGGAGCGAGGGTG 260
|||||
QY 262 CCGCAACAGCAGAGTCTCTGACCTGACGAGGAGCGGAGTGGAAACCGGGGAGAG 321
|||||
Db 261 CCGCAACAGCAGAGTCTCTGACCTGACGAGTGGAAACCGGGGAGAG 319
|||||
QY 322 AGCCAGGTGGCCAGGTACCAATGGCAATCATGTACCGGGGGTCTATGACTATCACT 381
|||||
Db 320 AGCCAGGTGGCCAGGTACCAATGGCAATCATGTACCGGGGGTCTATGACTATCACT 379
|||||
QY 382 GGCACATCTACATCTACAATGGACAGTACTGGGGGAGGACCGGGTCTGGAGACCTC 441
|||||
Db 380 GGCACATCTACATCTACAATGGACAGTACTGGGGGAGGACCGGGTCTGGAGACCTC 439
|||||
QY 442 CCAGCTACCCCGAAGCTTCATACCCCATTCGCGAGAGGGGAGCCTGGCCCTCCCGGG 501
|||||
Db 440 CCAGCTACCCCGAAGCTTCATACCCCATTCGCGAGAGGGGAGCCTGGCCCTCCCGGG 499
|||||
QY 502 CTCTCTACACCCACAGAGATGGCAAGGCTTGGCACCCTAGCGGAGACAGCACTGT 561
|||||
Db 500 CTCTCTACACCCACAGAGATGGCAAGGCTTGGCACCCTAGCGGAGACAGCACTGT 559
|||||
QY 562 GGTGCCACACCCCTC 575
|||||
Db 560 GGTGCCACACCCCTC 573
|||||

RESULT 8
BE740143
LOCUS BE740143
DEFINITION 601595085F1 NIH_MGC_9 Homo sapiens cdna clone IMAGE:3949148 5',
mRNA sequence.
ACCESSION BE740143
VERSION BE740143.1 GI:10154135


```

Db 241 CCACGGTACCAATGGCAATTCATGTCACCGCGGGTCTATGACTATCACTGGCAACATCTA 300
|||||
QY 393 CATCTACAATGACCACTACTGCGGGGACACCGGGTCTCTGGAGACCTCCAGCTACCC 452
|||||
Db 301 CATCTACAATGACCACTACTGCGGGGACACCGGGTCTCTGGAGACCTCCAGCTACCC 360
|||||
QY 453 CGAACCTCCATACCCATCCGCGGAGAGGGGACCCCTGCGCCCTCCCGGGCTCTCTACACC 512
|||||
Db 361 CGAACCTCCATACCCATCCGCGGAGAGGGGACCCCTGCGCCCTCCCGGGCTCTCTACACC 420
|||||
QY 513 CCACAGGAAGATGGCAAGGCTTGGACCTAGCGGAGACAGACACTGTGGTGCCACACC 572
|||||
Db 421 CCACAGGAAGATGGCAAGGCTTGGACCTAGCGGAGACAGACACTGTGGTGCCACACC 480
|||||
QY 573 CTC 575
|||
Db 481 CTC 483
|||

RESULT 10
LOCUS BE769498 900 bp mRNA linear EST 25-SEP-2001
DEFINITION 603059047F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5208394 5',
mRNA sequence.
ACCESSION BE769498
VERSION BE769498.1 GI:15761076
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 900)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tissue procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1522 row: n column: 11
High quality sequence stop: 854.
FEATURES
Location/Qualifiers
1..900
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5208394"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source:
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT 199 a 298 c 246 g 157 t
ORIGIN
Query Match 83.5%; Score 480.4; DB 13; Length 900;
Best Local Similarity 99.6%; Pred. No. 3e-108;
Matches 492; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 82 GGATCGCTGCTCAAGAGCGCTCGCAGGAGAGGGACCAATCTCTAGCTGGAAGCTGG 141
|||||

```

```

Db 269 GGATCGCTGCTCAAGAGCGCTCGCAGGAGAGGGACCAATCTCTAGCTGGAAGCTGG 328
|||||
QY 142 GAGCCTCCGAAGGCCCATCATCTCCCTGACTTCCCTGACTTGGTACAGCACTGCTACCCATTCT 201
|||||
Db 329 GAGCCTCCGAAGGCCCATCATCTCCCTGACTTCCCTGACTTGGTACAGCACTGCTACCCATTCT 388
|||||
QY 202 GGAGATGTTTCCCAAGTATCCACTGGGTCCCGCGCAGCCCCAGTGTGAGGCGAGGGGTG 261
|||||
Db 389 GGAGATGTTTCCCAAGTATCCACTGGGTCCCGCGCAGCCCCAGTGTGAGGCGAGGGGTG 448
|||||
QY 262 CGCAACACACAGAGTCTCTGGACCTGACACAGGAGCCCGAGTTGGAACCCCGGGGAGCAG 321
|||||
Db 449 CGCAACACACAGAGTCTCTGGACCTGACCA - GGAGCCCGAGTTGGAACCCCGGGGAGCAG 507
|||||
QY 322 AGCAGGTGGCCACAGGTACCAATGCAATTCATGTCAACCGGGGGTCTATGACTATCACT 381
|||||
Db 508 AGCAGGTGGCCACAGGTACCAATGCAATTCATGTCAACCGGGGGTCTATGACTATCACT 567
|||||
QY 382 GGCACATCTACATCTACAATGACACAGTACTGGGGGGACACCGGGTCTCTGGAGACCTC 441
|||||
Db 568 GGCACATCTACATCTACAATGACACAGTACTGGGGGGACACCGGGTCTCTGGAGACCTC 627
|||||
QY 442 CGAGCTACCCCGAAGCTCCATACCCATTCGGAAGAGGGGACCCCTGGCCCTCCCGGG 501
|||||
Db 628 CCAGCTTACCCCGAAGCTCCATACCCATTCGGAAGAGGGGACCCCTGGCCCTCCCGGG 687
|||||
QY 502 CTCTCTACACCCACAGGAGATGCAAGGCTTGGCAGCTAGCGGAGACAGACACTGT 561
|||||
Db 688 CTCTCTACAGCCACAGGAGATGCAAGGCTTGGCAGCTAGCGGAGACAGACACTGT 747
|||||
QY 562 GGTGTCACACCCCTC 575
|||||
Db 748 GGTGTCACACCCCTC 761
|||||

RESULT 11
LOCUS BE789430 749 bp mRNA linear EST 20-OCT-2000
DEFINITION 601482267F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3884920 5',
mRNA sequence.
ACCESSION BE789430
VERSION BE789430.1 GI:10210628
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 749)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9659 row: e column: 17
High quality sequence stop: 743.
FEATURES
Location/Qualifiers
1..749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3884920"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life

```

```

BASE COUNT      167 a      241 c      220 g      121 t
ORIGIN

Query Match      78.7%; Score 452.8; DB 12; Length 749;
Best Local Similarity 99.4%; Pred. No. 1.8e-101;
Matches 465; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 109 GGAGAGGACCCAA-TCTGTAGCTGGAAGCTGGAGAGCTCCGAAGGCCATCCATCACTT 167
|
|
|
Db 1 GGAGAGGACCCAAAGTCTGTAGCTGGAAGCTGGAGAGCTCCGAAGGCCATCCATCACTT 60
|
|
|
QY 168 CCTGACTTGGTACAGCAGCTCTACCCATTTCTGGAGATGTTCCCGAGTATCCACTGG 227
|
|
|
Db 61 CCTGACTTGGTACAGCAGCTCTACCCATTTCTGGAGATGTTCCCGAGTATCCACTGG 120
|
|
|
QY 228 GCTCCCGCAGCCAGTTTGGAGGAGGGGTGCGCAACAGCAGAGTCTCTGGAGCT 287
|
|
|
Db 121 GCTCCCGCAGCCAGTTTGGAGGAGGGGTGCGCAACAGCAGAGTCTCTGGAGCT 180
|
|
|
QY 288 GACCAGGAGCGCAGTTGGAAACCCGGGAGCAGAGCTGGCCCGTACCAATGG 347
|
|
|
Db 181 GACCAGGAGCGCAGTTGGAAACCCGGGAGCAGAGCTGGCCCGTACCAATGG 240
|
|
|
QY 348 CATTGATGTCACCGCGGTCTATGACTATGCTGGCAACATCTACATCTACATGGAGC 407
|
|
|
Db 241 CATTGATGTCACCGCGGTCTATGACTATGCTGGCAACATCTACATCTACATGGAGC 300
|
|
|
QY 408 AGTACTGGGGGACCAACCGGGTCTCTGGAGACCTCCAGCTACCCCGCACTCCATACCC 467
|
|
|
Db 301 AGTACTGGGGGACCAACCGGGTCTCTGGAGACCTCCAGCTACCCCGCACTCCATACCC 360
|
|
|
QY 468 CATTCCCGAAGAGGGGACCTGGCCCTCCCGGGGTCTCTACACCCCGCAAGAGATGG 527
|
|
|
Db 361 CATTCCCGAAGAGGGGACCTGGCCCTCCCGGGGTCTCTACACCCCGCAAGAGATGG 420
|
|
|
QY 528 CAAGCTTTGGCACTAGCGGAGACAGCACTGTGTGCCACACCCCTC 575
|
|
|
Db 421 CAAGCTTTGGCACTAGCGGAGACAGCACTGTGTGCCACACCCCTC 468

RESULT 12
BG331666      735 bp      mRNA      linear      EST 27-FEB-2001
LOCUS      602432673F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4550298 5',
DEFINITION      mRNA sequence.
ACCESSION      BG331666
VERSION      BG331666.1 GI:13138104
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 735)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1243 row: a column: 19
High quality sequence stop: 709.
Location/Qualifiers
1. 735
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4550298"

FEATURES
source
Location/Qualifiers
1. 735
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4550298"

/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT      165 a      232 c      219 g      119 t
ORIGIN

Query Match      77.4%; Score 445; DB 12; Length 735;
Best Local Similarity 99.6%; Pred. No. 1.5e-99;
Matches 467; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 108 GGGAGAGGACCCCAATCTGTAGCTGGAA-GCTGGAGCCTCCGAAGGCCATCCATCACT 166
|
|
|
Db 2 GGGAGAGGACCCCAATCTGTAGCTGGAAAGCTGGAGCCTCCGAAGGCCATCCATCACT 61
|
|
|
QY 167 TCCCTGACTTGGTACAGCACTGCTACCCATTTCTGGAGATGTTCCCGAGTATCCACTG 226
|
|
|
Db 62 TCCCTGACTTGGTACAGCACTGCTACCCATTTCTGGAGATGTTCCCGAGTATCCACTG 121
|
|
|
QY 227 GCCTCCCGCAGCCCGCAGTTTGGAGGAGGGGTGCGCAACAGCAGAGTCTCTGGAGC 286
|
|
|
Db 122 GCCTCCCGCAGCCCGCAGTTTGGAGGAGGGGTGCGCAACAGCAGAGTCTCTGGAGC 181
|
|
|
QY 287 TGACAGGAGGACCCAGTTGGAAACCCGGGAGCAGAGCCAGTGGCCCGTACCAATG 346
|
|
|
Db 182 TGACCA-GGAGCCGAGTTGGAACCCGGGAGCAGAGCCAGTGGCCCGTACCAATG 240
|
|
|
QY 347 GCATTGATGTCACCGCGGTCTATGACTATGCTGGCAACATCTACATCTACATGGAGC 406
|
|
|
Db 241 GCATTGATGTCACCGCGGTCTATGACTATGCTGGCAACATCTACATCTACATGGAGC 300
|
|
|
QY 407 CAGTACTGGGGGACCAACCGGGTCTCTGGAGACCTCCAGCTACCCCGCACTCCATACC 466
|
|
|
Db 301 CAGTACTGGGGGACCAACCGGGTCTCTGGAGACCTCCAGCTACCCCGCACTCCATACC 360
|
|
|
QY 467 CCATTCCCGAAGAGGGGACCTGGCCCTCCCGGGGTCTCTACACCCCGCAAGAGATG 526
|
|
|
Db 361 CCATTCCCGAAGAGGGGACCTGGCCCTCCCGGGGTCTCTACACCCCGCAAGAGATG 420
|
|
|
QY 527 GCAAGGCTTGGCACTAGCGGAGACAGCACTGTGTGCCACACCCCTC 575
|
|
|
Db 421 GCAAGGCTTGGCACTAGCGGAGACAGCACTGTGTGCCACACCCCTC 469

RESULT 13
BG331666      606 bp      mRNA      linear      EST 16-OCT-2001
LOCUS      603063085F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5212148 5',
DEFINITION      mRNA sequence.
ACCESSION      BG331666
VERSION      BG331666.1 GI:16168543
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 606)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

```


RESULT 15

BC403820
LOCUS BC403820 598 bp mRNA linear EST 12-MAR-2001
DEFINITION 602419660F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4526655 5',
mRNA sequence.

ACCESSION BG403820

VERSION BG403820.1 GI:13297268

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 598)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM10434 row: h column: 16

High quality sequence stop: 596.

FEATURES

source

1..598

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4526655"

/clone_lib="NIH_MGC_93"

/tissue_type="transitional cell papilloma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 1.7 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 133 a 193 c 171 g 101 t

ORIGIN

Query Match 75.1%; Score 432; DB 12; Length 598;

Best Local Similarity 99.6%; Pred. No. 2.3e-96;

Matches 454; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 121 AATCCTGTAGCTGAGCTGGAGCTCCGAGGCCATCCATACATCTCCCTGACTTGGTA 180

Db 1 AATCCTGTAGCTGAGCTGGAGCTCCGAGGCCATCCATACATCTCCCTGACTTGGTA 60

Qy 181 CAGCCACTGCTACCATTTCTGGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCC 240

Db 61 CAGCCACTGCTACCATTTCTGGAGATGTTTCCCGAGTATCCACTGGGCTCCCGCAGCC 120

Qy 241 CCAGTTTGGAGGAGGGGTGCCCAACAGCAGAGTCTCTGGACCTGACACAGGAGCGC 300

Db 121 CCAG--TTTGGAGGAGGGGTGCCCAACAGCAGAGTCTCTGGACCTGACACAGGAGCGC 179

Qy 301 CAGTTGGAACCCGGGAGCAGCAGCTGGCGCCACGCTACCAATGGCATTCATCTACC 360

Db 180 CAGTTGGAACCCGGGAGCAGCAGCTGGCGCCACGCTACCAATGGCATTCATCTACC 239

Qy 361 GCGGGGTCTATGACTATCACTGGCAACATCTACATCTACATGGACCCAGTACTGGGGGA 420

Db 240 GCGGGGTCTATGACTATCACTGGCAACATCTACATCTACATGGACCCAGTACTGGGGGA 299

Qy 421 CCACCGGGTCTTGAGACCTCCAGCTACCCCGAAGCTCCATACCCCATTCGCGAAGAG 480

Db 300 CCACCGGGTCTTGAGACCTCCAGCTACCCCGAAGCTCCATACCCCATTCGCGAAGAG 359

Qy 481 GGGGACCTGGCCCTCCCGGCTCTCTACACCCACGAGGAGATGCAAGGCT-TGGCA 539

Db 360 GGGACCCCTGCCCTCCCGGCTCTCTACACCCACGAGGATGGCAAGCTCTGGCA 419

Qy 540 CCTAGCGGAGACAGAGCACTGTGTGGCCACACCCCTC 575

Db 420 CCTAGCGGAGACAGAGCACTGTGTGGCCACACCCCTC 455

Search completed: April 15, 2003, 16:26:30

Job time : 669.156 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 13:11:20 ; Search time 18.528 seconds
(without alignments)
9517.464 Million cell updates/sec

Title: US-09-917-372-2_COPY_804_1378

Perfect score: 575

Sequence: 1 ctgtgtgagcagctccagg.....cactgtgtgtgcacacccctc 575

Scoring table: IDENTITY_NUC
Gapop.10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	35.2	6.1	111282	4	US-09-754-250-3
2	34.2	5.9	268	1	Sequence 3, Appli
3	34	5.9	3468	1	Sequence 16, Appli
4	34	5.9	3468	1	Sequence 2, Appli
5	34	5.9	3468	2	Sequence 2, Appli
6	34	5.9	3468	3	Sequence 2, Appli
7	34	5.9	3468	3	Sequence 2, Appli
8	34	5.9	3468	3	Sequence 2, Appli
9	34	5.9	3468	3	Sequence 2, Appli
10	33.2	5.8	4086	1	Sequence 1, Appli
11	33	5.7	1057	3	Sequence 1, Appli
12	33	5.7	1722	4	Sequence 18, Appli
13	33	5.7	2004	3	Sequence 14, Appli
14	32.8	5.7	376	4	Sequence 230, App
15	32.4	5.6	2473	4	Sequence 201, App
16	32.2	5.6	948	1	Sequence 4, Appli
17	32.2	5.6	948	2	Sequence 31, Appli
18	32.2	5.6	948	2	Sequence 31, Appli
19	32.2	5.6	948	2	Sequence 31, Appli
20	32.2	5.6	948	2	Sequence 31, Appli
21	32.2	5.6	948	2	Sequence 31, Appli
22	32.2	5.6	948	2	Sequence 26, Appli
23	32.2	5.6	948	3	Sequence 31, Appli
24	32.2	5.6	948	3	Sequence 31, Appli
25	32.2	5.6	948	3	Sequence 31, Appli
26	32.2	5.6	948	4	Sequence 31, Appli
27	32.2	5.6	948	4	Sequence 31, Appli

28	32.2	5.6	962	1	US-08-073-384C-11	Sequence 11, Appl
29	32.2	5.6	962	1	US-08-254-359A-11	Sequence 11, Appl
30	32.2	5.6	962	1	US-08-483-043-11	Sequence 11, Appl
31	32.2	5.6	962	1	US-08-481-238-11	Sequence 11, Appl
32	32.2	5.6	962	2	US-08-471-066B-11	Sequence 11, Appl
33	32.2	5.6	962	2	US-08-484-956-11	Sequence 11, Appl
34	32.2	5.6	962	2	US-08-757-653-11	Sequence 11, Appl
35	32.2	5.6	962	2	US-08-599-491-11	Sequence 11, Appl
36	32.2	5.6	962	2	US-08-756-386-11	Sequence 11, Appl
37	32.2	5.6	962	2	US-08-823-516-11	Sequence 11, Appl
38	32.2	5.6	962	3	US-08-682-853A-11	Sequence 11, Appl
39	32.2	5.6	962	3	US-08-759-038-11	Sequence 11, Appl
40	32.2	5.6	962	3	US-08-758-314-11	Sequence 11, Appl
41	32.2	5.6	962	4	US-09-350-309-11	Sequence 11, Appl
42	32.2	5.6	962	4	US-08-520-946-11	Sequence 11, Appl
43	32.2	5.6	963	2	US-08-757-653-162	Sequence 162, App
44	32.2	5.6	963	2	US-08-823-516-60	Sequence 60, Appl
45	32.2	5.6	963	3	US-08-759-038-101	Sequence 101, App

ALIGNMENTS

RESULT 1

US-09-754-250-3/c
; Sequence 3, Application US/09754250
; Patent No. 6376225
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001063
; CURRENT APPLICATION NUMBER: US/09/754,250
; CURRENT FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(111282)
; OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

Query Match 6.1%; Score 35.2; DB 4; Length 111282;
Best Local Similarity 50.6%; Pred. No. 3.4;
Matches 85; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
QY 2 TGGTGGAGCAGCTCCAGGCACTGCCAGTCCGACACCACTGCAACAAATCCATTAGAGC 61
DB 29256 TGGGAAAAGAGCCAGCGCTCCCGAGTCTCTTCTTGGCTTCAGGGGTGACTTTGAGC 29197
QY 62 CACTGCCCCAGAGATGTCAGGATCGCTCAAGAGGCGTCCGAGGAGGGACCCA 121
DB 29196 ACTCTCTCCAGCCAGCAGTCCCGCCAGCTCAGAGGACGACGACACTCCCTCT 29137
QY 122 ATCTGTAGCTGGAGCGCTCCGAGGCGCCATCCAGGCGCCATCACTATCTCC 169
DB 29136 CTCTTGCCTCCTCAAGTTGGAGCAACCTAGAGCGCTGGCCACACTTCC 29089

RESULT 2

US-08-039-137-16
; Sequence 16, Application US/08039137
; Patent No. 5759771
; GENERAL INFORMATION:
; APPLICANT: Tilanus J.G., Marcel
; TITLE OF INVENTION: Method of determining a Genotype by
; TITLE OF INVENTION: Comparing the Nucleotide Sequence of Members of a Gene
; Patent No. 5759771

RESULT 3
US-07-951-715A-2
Sequence 2, Application US/07951715A
; Patent No. 5625136
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalin M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Lumnis, Karen L.

RESULT 4
US-08-459-448A-2
; Sequence 2, Application US/08459448A
; Patent No. 5859336


```

CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1995
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3468 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..3468
OTHER INFORMATION: /product= "Full-length pure maize
optimized synthetic Bt"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3465
US-09-053-549-3

Query Match          5.9%; Score 34; DB 3; Length 3468;
Best Local Similarity 51.3%; Pred. No. 2.4;
Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 404 GACCAGTACTGGGGGACCAACCGGTCTCTGGAGACCTCCAGCTACCCCGGAACCTCCAT 463
      ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 937 GAGTACTACTGGAGCGGCCACAGATCATGCCAGCCCCGTGGCTTCACGGCGCCCGAG 996
      ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 464 ACCCATTTCCCAGAAGGGGACCCCTGGCCCTCCCGGGCTCTCTACACCCCCACCGAAG 523
      ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 997 TTCACCTTCCCCCTGTACGGCACCATGGCAACGCGCCCCCAGCAGCGCATCGTG GCC 1056
      ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 524 ATGGCAAGGCTTGGCACCTAGCGAGACAGCA 557
      ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1057 CAGCTGGGCGAGGGCGGTGTACCGCACCTGAGCA 1090
      ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-09-547-422-2
; Sequence 2, Application US/09547422
; Patent No. 6320100
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research, Inc
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/313,181
: FILING DATE: 07-OCT-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Reed, Janet E.
: REGISTRATION NUMBER: 36,252
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (215) 563-4100
: TELEFAX: (215) 563-4044
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4086 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: not relevant
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
:
: US-08-313-181-1
:
Query Match          5.8%; Score 33.2; DB 1;
Best Local Similarity 45.0%; Pred. No. 4.3;
Matches 125; Conservative 0; Mismatches 153;

Qy 228 GCTCCCGCGAGCCCGCAGTTTGGAGCGAGGGGTGCGGCACACGACG
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1724 GCGCACGCGCATCCGCTATATCGACGGCGCTCGAGGCTCTGCTGCC
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 288 GAGCAGGAGCGCCAGCTTGGAACCCCGGGCAGCAGACGACGGTGGG
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1784 CCCCCTGGCCGCGCAGCGCGCCTTCTATCGCCGGCGCGCGTGGTGC
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 348 CATTCTATGTCACGGCGGGGTCTATGACTATCTACTGTGCGAACATCTT
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1844 CGAGCACTACAGCGGGCGACTCCGACGCGTCCAGCCCGCGCTCCCA
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 408 AGTACTGGGGGGACACCGGGTCTCTGGAGACCTCCACGCTACCCG
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1904 GGTAGCCCGGACCCAGGAAGTGAGGAAGTTAGGGCGGCGCTCT
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 468 CATTCCCGAAGAGGGGACCCCTGGCCCTCCCGGGGTCT 505
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1964 CGTTTCGAGGGCGGGAGCTGCGCTTTCGGGAGGTTT 2001
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
US-08-313-181-1
; Sequence 1, Application US/08313181
; Patent No. 5681735
; GENERAL INFORMATION:
; APPLICANT: Emerson, Charles P.
; APPLICANT: Goldhamer, David J.
; TITLE OF INVENTION: Transcription Control Element for
; TITLE OF INVENTION: Increasing Gene Expression in Myoblasts
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street, Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```
; NAME/KEY: unsure
; LOCATION: (45)...(45)
; NAME/KEY: unsure
; LOCATION: (53)...(53)
; NAME/KEY: unsure
; LOCATION: (116)...(116)
; NAME/KEY: unsure
; LOCATION: (118)...(118)
US-09-188-930-18

Query Match
Best Local Similarity 5.7%; Score 33; DB 3; Length 1057;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 102 TCGCAGGAGGAGGACCAATCTGTAGCTGGAGCTGGAGCCCTCCGAGGCCCATCC 161
Db 353 TCGCAGCGTTGGATTCCGATCTCCAGCGGAGAGGTGCCCATTTACAGGCTCTTCC 412
QY 162 ATACTTCCCTGACTTGGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTATC 221
Db 413 CTAAGCCCTGGACAGATTGATCTCCAAGTACTCTCTCGCGGAGCTACACCTGTCAATCA 472
QY 222 CACTGGCTCCCGCAGCCCGCCAGTTTGTGGAGC 254
Db 473 CGCAAGGCTTTTGGAGGACCCGATACCTGTGGGGC 505

RESULT 12
US-09-142-623-14/C
; Sequence 14, Application US/09142623
; Patent No. 6337201
; GENERAL INFORMATION:
; APPLICANT: KOJI YANAI et al.
; TITLE OF INVENTION: -FRUCTOFURANOSIDASE AND ITS GENE, METHOD OF
; TITLE OF INVENTION: ISOLATING -FRUCTOFURANOSIDASE GENE, SYSTEM FOR PRODUCING
; TITLE OF INVENTION: -FRUCTOFURANOSIDASE, AND -FRUCTOFURANOSIDASE VARIANT
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESS: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,623
; FILING DATE: September 10, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 98-0989*/LC(WMC)/144
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1722 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: Double stranded
; TOPOLOGY: Linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Microorganism: Scopulariopsis brevicaulis IFO4843
; FEATURE:
; NAME/KEY: mat peptide
```

```
; LOCATION: 1...1722
; IDENTIFICATION METHOD: E
US-09-142-623-14

Query Match
Best Local Similarity 5.7%; Score 33; DB 4; Length 1722;
Matches 69; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 345 TGCATTATGTCACCGGGGCTCTATGACTATCATCTGCGAATCATCTACATCTACATGG 404
Db 838 TGCAGAAGATTTTACCGTGGGGTTGTAGCCATCATCTGTCAGACGACAAATTTGTATGA 779
QY 405 ACCAGTACTGGGGGGACCCGCGGTCTCTGGAGACCTCCAGCTACCCCGCAACCTCCATA 464
Db 778 CCTCGAAGTTGTAGCCCGCCAGCCAGTCACCGCTGCCCGAGTGCAGTTCCTT 719
QY 465 CCCCATTC 473
Db 718 CCTCGTTCC 710

RESULT 13
US-09-188-930-230
; Sequence 230, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011cl
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 230
; LENGTH: 2004
; TYPE: DNA
; ORGANISM: Rat
US-09-188-930-230

Query Match
Best Local Similarity 5.7%; Score 33; DB 3; Length 2004;
Matches 78; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 102 TCGCAGGAGGAGGAGCCCAATCTGTAGCTGGAAAGCTGGAGCCCTCCGAGGCCCATCC 161
Db 287 TCGCAGCGGTTGGGATTCGGATCTGCAGCGGGAAGAGGTGCCCATTTACAGGCTCTTCC 346
QY 162 ATACTTCCCTGACTTGGTACAGCCACTGCTACCCATTTCTGGAGATGTTTCCCGAGTATC 221
Db 347 CTAAGCCCTGGGACAGATTGATCTCCAAGTACTCTCTCGCGGAGCTACACCTGTCAATCA 406
QY 222 CACTGGCTCCCGCAGCCCGCCAGTTTGTGGAGC 254
Db 407 CGCAAGGCTTTTGGAGGACCCGATACCTGTGGGGC 439

RESULT 14
US-09-149-476-201/c
; Sequence 201, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
```

us-09-917-372-2_copy_804_1378.rni

[illegible]

EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 5.7% Score 32.8; DB 4; Length 376;

Best Local Similarity 47.5%; Pred. No. 2.6;

Matches 94; Conservative 1; Mismatches 103; Indels 0; Gaps 0;

Qy 111 AGAGGGACCAATCCTGTAGTGGAGGCTCCGAGGCCATCCAGGCCCATCTTCCC 170
Db 354 AACGAGCCAGATGCTTTAATGGGCTGGCTCCAGGCCCATCGGCTCATCAGATCCCTT 295
Qy 171 TGACTTGGTACAGCACTGTACCAATTCCTGGAGATGTTTCCCAAGTATCCACTGGGCT 230
Db 294 AATGAGGCCCCAGAGCTTGGCAGCCTTTGGTTGAGCTCAGCACTCCACAGGCGGTGGCT 235
Qy 231 CCCCGCAGCCCACTTTGGAGGCGGCTCCCAACAGCAGAGTCTCTGGACCTGAC 290
Db 234 CACAGAGGCGCTGCTGATGGGCTCTGGCTGGGGAGTAGCARGGTCTCTCAGTCCAGATC 175
Qy 291 CAGGAGCCCGAGTTGGA 308
Db 174 TTCAGAGACCTTCTAGGA 157

RESULT 15

US-08-794-494-4/c

; Sequence 4, Application US/08794494

; Patent No. 5981833

; GENERAL INFORMATION:

; APPLICANT: Wise, Roger P.

; APPLICANT: Schnable, Patrick S.

TITLE OF INVENTION: NUCLEAR RESTORER GENES FOR HYBRID SEED PRODUCTION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leydig, Voit & Mayer, Ltd.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/794,494
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 8-346611
FILING DATE: 29-NOV-1994
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Larcher, Carol
REGISTRATION NUMBER: 35243
REFERENCE/DOCKET NUMBER: 74959
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5600
TELEFAX: (312) 616-5700
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2473 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-794-494-4

Query Match 5.6% Score 32.4; DB 2; Length 2473;

Best Local Similarity 55.3%; Pred. No. 6.1;

Matches 63; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 426 GGGTCTTGGAGACCTCCAGCTACCCCGAAGCCATTCATCCCATTCCTCCGAGAGGGGA 485
Db 2031 GCGACCCGGGCATGCCGCCGAAGCCGCTGAAGCCCGGGGAGCCGCCCGAAGAGGGCA 1972
Qy 486 CCCTGGCCCTCCCGGGCTCTTACACCCACCAGAGATGGCAAGGCTTGGCA 539
Db 1971 TGCTGCGGCTCGCGGCAAGCTGAACCCGGGTAGGCTGCTGCCGCTGCCGGCA 1918

Search completed: April 15, 2003, 16:29:52

Job time : 95.528 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 15, 2003, 15:24:11 ; Search time 40.0647 Seconds
(without alignments)
12588.918 Million cell updates/sec

Title: US-09-917-372-2_COPY_804_1378

Perfect score: 575

Sequence: 1 ctgtgtgagcagctccagg.....cactgtgtgccacacctc 575

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	575	100.0	1982	10	US-09-907-372-2
2	561.2	97.6	574	10	US-09-907-372-8
3	494	85.9	2136	9	US-09-954-531-1348
4	494	85.9	2136	10	US-09-962-436-262
5	494	85.9	2136	10	US-09-880-107-2135
c 6	452	78.6	651	10	US-09-907-372-7
c 7	425	73.9	425	10	US-09-907-372-9
c 8	253.8	44.1	281	10	US-09-867-701-5601
c 9	251.2	43.7	289	10	US-09-867-701-5666
10	240.4	41.8	371	10	US-09-907-372-16
11	198.4	34.5	219	10	US-09-907-372-10
12	172	29.9	279	10	US-09-907-372-11
13	144.2	25.1	206	10	US-09-907-372-13
c 14	101.8	17.7	862	10	US-09-907-372-12
15	61.2	10.6	481	10	US-09-867-701-4881
16	39.2	6.8	548	10	US-09-907-372-14
c 17	36.4	6.3	2192	10	US-09-833-381-906
18	36	6.3	471	10	US-09-907-372-15
19	35.4	6.2	809	10	US-09-844-864-22

c 20	35.2	6.1	111282	12	US-10-094-989-3	Sequence 3, Appli
c 21	34.6	6.0	2930	10	US-09-745-763-198	Sequence 198, App
c 22	34.4	6.0	2768	10	US-09-905-983-4	Sequence 4, Appli
c 23	34.4	6.0	2768	10	US-09-905-983-6	Sequence 6, Appli
c 24	34	5.9	3468	9	US-09-988-462-2	Sequence 2, Appli
c 25	34	5.9	6788	9	US-09-870-759-101	Sequence 101, App
c 26	33.6	5.8	33675	10	US-09-921-992-2	Sequence 2, Appli
c 27	33.2	5.8	2855	9	US-10-125-540-594	Sequence 594, App
c 28	33.2	5.8	2855	10	US-09-764-870-594	Sequence 594, App
c 29	33.2	5.8	15061	9	US-10-092-154-991	Sequence 991, App
c 30	33.2	5.8	15061	10	US-09-764-847-991	Sequence 991, App
c 31	33	5.7	1057	9	US-10-152-661-18	Sequence 18, Appli
c 32	33	5.7	1057	9	US-09-866-050A-18	Sequence 18, Appli
c 33	33	5.7	1722	9	US-09-990-385-14	Sequence 14, Appli
c 34	33	5.7	1890	9	US-10-152-661-447	Sequence 447, App
c 35	33	5.7	1890	9	US-09-866-050A-447	Sequence 447, App
c 36	33	5.7	2004	9	US-10-152-661-230	Sequence 230, App
c 37	33	5.7	2004	9	US-09-866-050A-230	Sequence 230, App
c 38	32.8	5.7	376	9	US-09-809-391-201	Sequence 201, App
c 39	32.2	5.6	595	10	US-09-878-574-4557	Sequence 4557, Ap
c 40	32.2	5.6	785	9	US-09-902-941-1868	Sequence 1868, Ap
c 41	32.2	5.6	785	9	US-09-849-626-1868	Sequence 1868, Ap
c 42	32.2	5.6	785	9	US-10-017-754-1868	Sequence 1868, Ap
c 43	32.2	5.6	948	9	US-10-033-297-26	Sequence 26, Appli
c 44	32.2	5.6	948	9	US-10-081-806-31	Sequence 31, Appli
c 45	32.2	5.6	948	9	US-10-074-328-31	Sequence 31, Appli

ALIGNMENTS

RESULT 1

US-09-907-372-2
; Sequence 2, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1982
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7497867CB1
US-09-907-372-2

Query Match	100.0%	Score	575;	DB	10;	Length	1982;
Best Local Similarity	100.0%	Pred. No.	2.9e-167;				
Matches	575;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	CTGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAAATCCATTAGAG	60				
Db	804	CTGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACAACTGCAAAAATCCATTAGAG	863				
Qy	61	CCACTGCCCCAGAGATGTCAGGATCGTGTCTCAAGAGCGCTCCGACGAGAGAGGAGCC	120				
Db	864	CCACTGCCCCAGAGATGTCAGGATCGTGTCTCAAGAGCGCTCCGACGAGAGAGGAGCC	923				
Qy	121	AATCTGTAGCTGGAAGCTGGGAGCTCCGAAGGCCCATCTCCATCTCCCTGACTTGGTA	180				
Db	924	AATCTGTAGCTGGAAGCTGGGAGCTCCGAAGGCCCATCTCCATCTCCCTGACTTGGTA	983				
Qy	181	CAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTATCCACTGGGCTCCCCCAGCC	240				
Db	984	CAGCCACTGCTACCCATTTCTGGAGATGTTTCCCCAGTATCCACTGGGCTCCCCCAGCC	1043				

Db 301 CAGTTGGAAACCCGGGAGCAGACCCAGGTGGCCACGGTACCAATGGCATTATGTACAC 360
QY 361 GCGGGGTCTATGACTATACACTGGCAACATCTACATCTACAATGGACCACTACTGGGGGA 420
Db 361 GCGGGGTCTATGACTATACACTGGCAACATCTACATCTACAATGGACCACTACTGGGGGA 420
QY 421 CCACCGGCTCTGAGACCTCCAGCTACCCCGGAACCTTCCATACCCATTCCTCCGAAGAG 480
Db 421 CCACCGGCTCTGAGACCTCCAGCTACCCCGGAACCTTCCATACCCATTCCTCCGAAGAG 480
QY 481 GGGGACCTTGGCCCTCCCGGGCTCTTAAACCCACCCAGGAGATGGCAAGCTTGGGAC 540
Db 481 GGGGACCTTGGCCCTCCCGGGCTCTTAAACCCACCCAGGAGATGGCAAGCTTGGGAC 540
QY 541 CTAGCGGAGACAGACACTGTGGTGC 566
Db 541 CTAGCGGAGACAGACACTGTGGTGC 566
RESULT 3
US-09-954-531-1348
; Sequence 1348, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1348
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-1348
Query Match 85.9%; Score 494; DB 9; Length 2136;
Best Local Similarity 100.0%; Pred. No. 2.4e-142;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 82 GGATCGTGTCTCAAGAGCGCTCCGAGGAGGAGGAGCCCAATCTCTAGCTGGAAGCTGG 141
Db 943 GGATCGTGTCTCAAGAGCGCTCCGAGGAGGAGGAGCCCAATCTCTAGCTGGAAGCTGG 1002
QY 142 GAGCTCCGAGGCGCCATCCATCTCCCTGACTTGGTACAGCCACTGCTACCATTTCT 201
Db 1003 GAGCTCCGAGGCGCCATCCATCTCCCTGACTTGGTACAGCCACTGCTACCATTTCT 1062
QY 202 GGAGATGTTTCCCAAGTATCCACTGGGCTCCCGACGCCAGTTTGGAGGAGGAGG 261
Db 1063 GGAGATGTTTCCCAAGTATCCACTGGGCTCCCGACGCCAGTTTGGAGGAGGAGG 1122
QY 262 CGGCAACAGCAGAGTCTCTTGACCTGACACGAGGAGCGGAGTTGGAAACCCGGGAGCAG 321
Db 1123 CGGCAACAGCAGAGTCTCTTGACCTGACACGAGGAGCGGAGTTGGAAACCCGGGAGCAG 1182
QY 322 AGCCAGGTGGCCACCGGTACCAATGGGATTCATGTACCGCGGGTCTATGACTATCACT 381
Db 1183 AGCCAGGTGGCCACCGGTACCAATGGGATTCATGTACCGCGGGTCTATGACTATCACT 1242
QY 382 GGCACATCTACATCTACAATGGACCTAGTGGGGGAGCCACCGGGTCTGTGGAGACCTC 441

QY 241 CCAGTTTGGAGCAGGAGTGGCGCAACAGAGAGTCTCTTGACCTGACCAAGGAGCGG 300
Db 1044 CCAGTTTGGAGCAGGAGTGGCGCAACAGAGAGTCTCTTGACCTGACCAAGGAGCGG 1103
QY 301 CAGTTGAAACCCGGGAGCAGACCCAGGTGGCCACGGTACCAATGGCATTATGTACAC 360
Db 1104 CAGTTGAAACCCGGGAGCAGACCCAGGTGGCCACGGTACCAATGGCATTATGTACAC 1163
QY 361 GCGGGGTCTATGACTATACACTGGCAACATCTACATCTACAATGGACCACTACTGGGGGA 420
Db 1164 GCGGGGTCTATGACTATACACTGGCAACATCTACATCTACAATGGACCACTACTGGGGGA 1223
QY 421 CCACCGGCTCTGAGACCTCCAGCTACCCCGGAACCTTCCATACCCATTCCTCCGAAGAG 480
Db 1224 CCACCGGCTCTGAGACCTCCAGCTACCCCGGAACCTTCCATACCCATTCCTCCGAAGAG 1283
QY 481 GGGGACCTTGGCCCTCCCGGGCTCTTAAACCCACCCAGGAGATGGCAAGCTTGGGAC 540
Db 1284 GGGGACCTTGGCCCTCCCGGGCTCTTAAACCCACCCAGGAGATGGCAAGCTTGGGAC 1343
QY 541 CTAGCGGAGACAGACACTGTGGTGC 566
Db 1344 CTAGCGGAGACAGACACTGTGGTGC 566
RESULT 2
US-09-907-372-8
; Sequence 8, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 8234468H1
US-09-907-372-8
Query Match 97.6%; Score 561.2; DB 10; Length 574;
Best Local Similarity 99.5%; Pred. No. 3.9e-163;
Matches 563; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CTGGTGGAGGAGCTCCAGGACTGCGGAGTCCGACACACCTGCAAAATCCATTAGAG 60
Db 1 CTGGTGGAGGAGCTCCAGGACTGCGGAGTCCGACACACCTGCAAAATCCATTAGAG 60
QY 61 CCATGCCCCCAGAGATGTAGAGTCTGCTCAAGAGGCGTCCGAGGAGGAGGAGCC 120
Db 61 CCATGCCCCCAGAGATGTAGAGTCTGCTCAAGAGGCGTCCGAGGAGGAGGAGCC 120
QY 121 AATCTCTAGTGAAGCTGGAGCTCCGAGGAGTCTCCAGTATCCACTTGGGCTCCCGGAGCC 240
Db 121 AATCTCTAGTGAAGCTGGAGCTCCGAGGAGTCTCCAGTATCCACTTGGGCTCCCGGAGCC 240
QY 181 CAGCCACTGCTACCCATTCTGGAGATGTTTCCCGAGTATCCACTTGGGCTCCCGGAGCC 300
Db 181 CAGCCACTGCTACCCATTCTGGAGATGTTTCCCGAGTATCCACTTGGGCTCCCGGAGCC 300
QY 241 CCAGTTTGGAGGAGGAGTGGCGCAACAGCAGAGTCTCTGGACCTGACCAAGGAGCCG 360
Db 241 CCAGTTTGGAGGAGGAGTGGCGCAACAGCAGAGTCTCTGGACCTGACCAAGGAGCCG 360
QY 301 CAGTTGAAACCCGGGAGCAGACCCAGGTGGCCACGGTACCAATGGCATTATGTACAC 360

Db	1243	GGCAACATCTACATCTACAACTGGACAGTACCTGGGGGGAACACACCGGTCTCTGGAGACCTC	1302
Qy	442	CCAGCTACCCCGGAACCTCCATACCCCAATCCCGAAGAGGGGACCTTGCGCCCTCCCGGG	501
Db	1303	CCAGCTACCCCGGAACCTCCATACCCCAATCCCGAAGAGGGGACCTTGCGCCCTCCCGGG	1362
Qy	502	CTCTCTACACCCACACAGAAGATGCAAGGCTTGCGACCTAGCGGAGACAGACACTGT	561
Db	1363	CTCTCTACACCCACCGAGAAGATGCAAGGCTTGCGACCTAGCGGAGACAGACACTGT	1422
Qy	562	GGTGCCACACCCCTC	575
Db	1423	GGTGCCACACCCCTC	1436

```

RESULT 4
US-09-962-436-262
; Sequence 262, Application US/09962436
; Patent NO. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 262
; LENGTH: 2136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-262

```

Query Match	85.98;	Score 494;	DB 10;	Length 2136;
Best Local Similarity	100.0%;	Prod. No. 2.4e-142;		
Matches 494;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;
Qy 82	GGATCGCTGCTCAAGAGCGCTCCGAGGGAGGAGGACCAATCCTGTAGCTGGAGACTGG 141			
Db 943	GGATCGCTGCTCAAGAGCGCTCCGAGGGAGGAGGACCAATCCTGTAGCTGGAGACTGG 1002			
Qy 142	GAGCTCTCGAAGGCCCATCCATCTTCCTGACTTGGTACAGCCACTGCTACCCATTTCT 201			
Db 1003	GAGCTCTCGAAGGCCCATCCATCTTCCTGACTTGGTACAGCCACTGCTACCCATTTCT 1062			
Qy 202	GGAGATGTTTCCCCAGTATCCACTGGGCTCCCGCAGGCCCATGTTTGGAGGCAGGGGTG 261			
Db 1063	GGAGATGTTTCCCCAGTATCCACTGGGCTCCCGCAGGCCCATGTTTGGAGGCAGGGGTG 1122			
Qy 262	CCGCAACAGCAGAGTCTCTGGAACCTGACAGGAGCGGCAGTTTGGAAACCCGGGGAGCAG 321			
Db 1123	CCGCAACAGCAGAGTCTCTGGAACCTGACAGGAGCGGCAGTTTGGAAACCCGGGGAGCAG 1182			
Qy 322	AGCCAGGTGGCCCGACGGTACCAATGGCAATTCATGTACCGGGGTCTATGACTATCACT 381			
Db 1183	AGCCAGGTGGCCCGACGGTACCAATGGCAATTCATGTACCGGGGTCTATGACTATCACT 1242			
Qy 382	GGCAACATCTACATCTACAATGGACCACTACTGGGGGAGCACCCGGTCTCTGGAGACCTC 441			
Db 1243	GGCAACATCTACATCTACAATGGACCACTACTGGGGGAGCACCCGGTCTCTGGAGACCTC 1302			
Qy 442	CCAGCTACCCCGGCAACCTCCATACCCCATTCCTCCGAAGAGGGGAGCCCTGGCCCTCCCGGG 501			
Db 1303	CCAGCTACCCCGGCAACCTCCATACCCCATTCCTCCGAAGAGGGGAGCCCTGGCCCTCCCGGG 1362			
Qy 502	CTCTCTACACCCCAACAGGAAGATGGCAAGCTTGGCACTAGCGGAGACAGAGCACTGT 561			

Db	1363	CTCTCTACACCCACCACGAGNAGTGCACAGGCTTGGCACCTTAGCGGACACAGACGACTGT	1422
Qy	562	GGTGCCACACCCCTC	575
Db	1423	GGTGCCACACCCCTC	1436

RESULT 5

US-09-880-107-2135

; Sequence 2135, Application US/09880107

; Patent No. US20020142981A1

; GENERAL INFORMATION:

; APPLICANT: Horne, Darci T.

; APPLICANT: Vockley, Joseph G.

; APPLICANT: Scherif, Uwe

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

; FILE REFERENCE: 44921-5028-WO

; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/211,379

; PRIOR FILING DATE: 2000-06-14

; PRIOR APPLICATION NUMBER: US 60/237,054

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 3950

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2135

; LENGTH: 2136

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L04270

US-09-880-107-2135

Query Match	85.94;	Score	494;	DB	10;	Length	2136;
Best Local Similarity	100.0%;	Pred. No.	2.4e-142;				
Matches	494;	Conservative	0;	Mismatches	0;	Indels	0;
							Gaps
QY	82	GGATCGCTGCTCAAGAGCGCTCCGAGGAGGAGGACCAATCTGCTAGCTGGAGACTGG	141				
Db	943	GGATCGCTGCTCAAGAGCGCTCCGAGGAGGAGGACCAATCTGCTAGCTGGAGACTGG	1002				
QY	142	GAGCTCCGAAAGCGCCATCCATATCTCCCTGACTTGGTACAGCCACTGCTACCCATTCT	201				
Db	1003	GAGCTCCGAGGCCATCCATATCTCCCTGACTTGGTACAGCCACTGCTACCCATTCT	1062				
QY	202	GGAGATGTTTCCCCAGTATTCACACTGGGCTCCCGAGGCCCACTGTTTGGAGCGAGGGTG	261				
Db	1063	GGAGATGTTTCCCCAGTATTCACACTGGGCTCCCGAGGCCCACTGTTTGGAGCGAGGGTG	1122				
QY	262	CGGCAACAGCAGAGTCCTGCGACCTGACAGGAGCGCAGTTCGAAACCCGGGGAGCAG	321				
Db	1123	CGGCAACAGCAGAGTCCTGCGACCTGACAGGAGCGCAGTTCGAAACCCGGGGAGCAG	1182				
QY	322	AGCCAGGTGGCCACGCGTACCAATGGCAATTGTCACCGCGGCTCTATGACTATCACT	381				
Db	1183	AGCCAGGTGGCCACGCGTACCAATGGCAATTGTCACCGCGGCTCTATGACTATCACT	1242				
QY	382	GGCAACATCTACATCTACAATGGACCACTAGTGGGGGAGACACCGGGTCTCTGGAGACCTC	441				
Db	1243	GGCAACATCTACATCTACAATGGACCACTAGTGGGGGAGACACCGGGTCTCTGGAGACCTC	1302				
QY	442	CCAGTACCCCGGAACCTCCATACCCCAATCCCGAAGAGGGGACCCCTGGCCCTCCCGGG	501				
Db	1303	CCAGTACCCCGGAACCTCCATACCCCAATCCCGAAGAGGGGACCCCTGGCCCTCCCGGG	1362				
QY	502	CTCTCTACACCCCAACAGGAAGTGGCAAGGCTTGGCACTCTAGCGGAGACAGACACTGT	561				
Db	1363	CTCTCTACACCCCAACAGGAAGTGGCAAGGCTTGGCACTCTAGCGGAGACAGACACTGT	1422				
QY	562	GGTGGCAACACCTC	575				
Db	1423	GGTGGCAACACCTC	1436				

```

; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7716340H1
US-09-907-372-9

Query Match      73.9%; Score 425; DB 10; Length 425;
Best Local Similarity 100.0%; Pred. No. 2.9e-121; Indels 0; Gaps 0;
Matches 425; Conservative 0; Mismatches 0;

QY 50 ATCCATTAGAGCCACTGCCCCAGAGATGTGAGGATCGCTCAAGAGGCGTCCGAGG 109
DB 425 ATCCATTAGAGCCACTGCCCCAGAGATGTGAGGATCGCTCAAGAGGCGTCCGAGG 366
QY 110 GAGAGGACCCAAATCCTGTAGCTGGAGCTGGAGCCCTCCGAAGGCCATCCATCTCC 169
DB 365 GAGAGGACCCAAATCCTGTAGCTGGAGCTGGAGCCCTCCGAAGGCCATCCATCTCC 306
QY 170 CTGACTGTGTACAGCCACTGTACCCATTTCTGGAGATGTTTCCCCAGTATCCACTGGGC 229
DB 305 CTGACTGTGTACAGCCACTGTACCCATTTCTGGAGATGTTTCCCCAGTATCCACTGGGC 246
QY 230 TCCCGGAGCCCGAGTTTGGAGGAGGGGTGCGCCAAACAGCAGAGTCTCTGGACCTGA 289
DB 245 TCCCGGAGCCCGAGTTTGGAGGAGGGGTGCGCCAAACAGCAGAGTCTCTGGACCTGA 186
QY 290 CCAGGAGCCCGAGTTGGAACCCGGGAGCAGAGCAGGTGGCCCAACGTTACCAATGSCA 349
DB 185 CCAGGAGCCCGAGTTGGAACCCGGGAGCAGAGCAGGTGGCCCAACGTTACCAATGSCA 126
QY 350 TTCACTGTACCGGGGGTCTATGACTATACCTGGAACATCTACATCTACATGAGACCA 409
DB 125 TTCACTGTACCGGGGGTCTATGACTATACCTGGAACATCTACATCTACATGAGACCA 66
QY 410 TACTGGGGGACACCGGGTCTGGAGACTCCAGTACCCCGAAGCTCCATACCCCA 469
DB 65 TACTGGGGGACACCGGGTCTGGAGACTCCAGTACCCCGAAGCTCCATACCCCA 6

QY 470 TTCCC 474
DB 5 TTCCC 1

RESULT 8
US-09-867-701-5601
; Sequence 5601, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5601
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-5601

Query Match      44.1%; Score 253.8; DB 10; Length 281;
Best Local Similarity 97.1%; Pred. No. 1.2e-68; Indels 1; Gaps 1;
Matches 269; Conservative 0; Mismatches 7;

QY 193 CCCATTTCTGGAGATGTTTCCCACTATCCACTGGGTCCCGCAGCCC-CAGTTTGGGA 251
```

```

; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7716364H1
US-09-907-372-7

Query Match      78.6%; Score 452; DB 10; Length 651;
Best Local Similarity 99.6%; Pred. No. 1.6e-129; Indels 2; Gaps 2;
Matches 474; Conservative 0; Mismatches 0;

QY 1 CTGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACCACTCCAAAATCCATTAGAG 60
DB 476 CTGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACCACTCCAAAATCCATTAGAG 417
QY 61 CCAGTGTCCCGGAGAGATGTGAGTGTCTCAAGAGGCTCGCAGGAGAGGAGCC 120
DB 416 CCAGTGTCCCGGAGAGATGTGAGTGTCTCAAGAGGCTCGCAGGAGAGGAGCC 357
QY 121 ATCTGTGTAGTGGAGCTGGAGCTCCGAAGGCCATCCATCTTCCCTGACTTGSTA 180
DB 356 ATCTGTGTAGTGGAGCTGGAGCTCCGAAGGCCATCCATCTTCCCTGACTTGSTA 297
QY 181 CAGCCACTGTACCCATTTCTGGAGATGTTCCCAAGTATCCACTGGGTCCCGCAGCC 240
DB 296 CAGCCACTGTACCCATTTCTGGAGATGTTCCCAAGTATCCACTGGGTCCCGCAGCC 237
QY 241 CCAGTTTTGGAGCAGGGTGGCCGACAGAGTCTCTGAGCTGACCAAGGAGCGG 300
DB 236 CCAGTTTTGGAGCAGGGTGGCCGACAGAGTCTCTGAGCTGACCAAGGAGCGG 177
QY 301 CAGTTGGAACCCGGGAGCAGAGCCAGGTGGCCCAAGTACCAATGGCATTCATGTACC 360
DB 176 CAGTTGGAACCCGGGAGCAGAGCCAGGTGGCCCAAGTACCAATGGCATTCATGTACC 117
QY 361 GCGGGTCTATGACTATCACTGCAACATCTACATCTCAATGGACCACTACTGGGGGA 420
DB 116 GCGGGTCTATGACTATCACTGCAACATCTACATCTCAATGGACCACTACTGGGGGA 57
QY 421 CCACCGGGTCTGAGACCTCC-AGCTACCCCGGAGACCTCC-ATACCCCATTC 474
DB 56 CCACCGGGTCTGAGACCTCCAGTACCTAGTACCTCCCGGAGACCTCCATACCCCATTC 1

RESULT 7
US-09-907-372-9/c
; Sequence 9, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 7716364H1
US-09-907-372-7

Query Match      78.6%; Score 452; DB 10; Length 651;
Best Local Similarity 99.6%; Pred. No. 1.6e-129; Indels 2; Gaps 2;
Matches 474; Conservative 0; Mismatches 0;

QY 1 CTGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACCACTCCAAAATCCATTAGAG 60
DB 476 CTGGTGGAGGAGCTCCAGGCACTGCCAGTCCGACACCACTCCAAAATCCATTAGAG 417
QY 61 CCAGTGTCCCGGAGAGATGTGAGTGTCTCAAGAGGCTCGCAGGAGAGGAGCC 120
DB 416 CCAGTGTCCCGGAGAGATGTGAGTGTCTCAAGAGGCTCGCAGGAGAGGAGCC 357
QY 121 ATCTGTGTAGTGGAGCTGGAGCTCCGAAGGCCATCCATCTTCCCTGACTTGSTA 180
DB 356 ATCTGTGTAGTGGAGCTGGAGCTCCGAAGGCCATCCATCTTCCCTGACTTGSTA 297
QY 181 CAGCCACTGTACCCATTTCTGGAGATGTTCCCAAGTATCCACTGGGTCCCGCAGCC 240
DB 296 CAGCCACTGTACCCATTTCTGGAGATGTTCCCAAGTATCCACTGGGTCCCGCAGCC 237
QY 241 CCAGTTTTGGAGCAGGGTGGCCGACAGAGTCTCTGAGCTGACCAAGGAGCGG 300
DB 236 CCAGTTTTGGAGCAGGGTGGCCGACAGAGTCTCTGAGCTGACCAAGGAGCGG 177
QY 301 CAGTTGGAACCCGGGAGCAGAGCCAGGTGGCCCAAGTACCAATGGCATTCATGTACC 360
DB 176 CAGTTGGAACCCGGGAGCAGAGCCAGGTGGCCCAAGTACCAATGGCATTCATGTACC 117
QY 361 GCGGGTCTATGACTATCACTGCAACATCTACATCTCAATGGACCACTACTGGGGGA 420
DB 116 GCGGGTCTATGACTATCACTGCAACATCTACATCTCAATGGACCACTACTGGGGGA 57
QY 421 CCACCGGGTCTGAGACCTCC-AGCTACCCCGGAGACCTCC-ATACCCCATTC 474
DB 56 CCACCGGGTCTGAGACCTCCAGTACCTAGTACCTCCCGGAGACCTCCATACCCCATTC 1
```

Db 5 CCACCTTTCTGGAGATCTTCCCGAGTATCCACTGGCTCCCGCAGCCGCGAGTTTGGG 64
QY 252 GGCAGGGTGGCCGACACAGAGTCTCTGGACCTGACGAGGAGCCGAGTTGGAAC 311
Db 65 GGCAGGGTGGCCGACACAGAGTCTCTGGACCTGACGAGGAGCCGCGTTGGAAC 124
QY 312 CGGGAGCAGAGCAGGTGCGCCACAGTACCAATGTCATTGTCACCGCGGGTCTAT 371
Db 125 CGGGAGCAGACTCAGTGGCCACAGTACCAATGTCATTGTCACCGCGGGTCTAT 184
QY 372 GACTATCTGCGCAACATCTACATCTACATGACAGTACTGCGGGGACCCAGCGGTCC 431
Db. 185 GACTATCTGCGCAACATCTACATCTACATGACAGTACTGCGGGGACCCAGCGGTCA 244
QY 432 TGGAGACCTCCCGAGTACCCCGCAACCTCCATACCCC 468
Db 245 TGGAGACCTCCCGAGTACCCCGCAACCTCCATTCCCC 281

RESULT 9
US-09-867-701-5666/c
; Sequence 5666, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5666
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-5666

Query Match 43.7%; Score 251.2; DB 10; Length 289;
Best Local Similarity 98.5%; Pred. No. 7.8e-68;
Matches 264; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 204 AGATGTTTCCCGAGTATCCACTGGCTCCCGCAGCCCGAGTTTGGAGGCGGTGCC 263
Db 289 AGATGTTTCCCGAGTATCCACTGGCTCCCGCAGCCCGAGTTTGGAGGCGGTGCC 230
QY 264 GCAACAGCAGA-GTCTCTGAGCTGACACGAGGAGCGCGAGTTGGAACCGGGGAGCAGA 322
Db 229 GCAACAGCAGATCTCTGAGCTGACACGAGGAGCGCGGTTGGAACCGGGGAGCAGA 170
QY 323 GCAGGTGGCCCGAGTACCAATGGAATTCATGTCACCGCGGGTCTATGACTATCACTG 382
Db 169 GCAGGTGGCCCGAGTACCAATGGAATTCATGTCACCGCGGGTCTATGACTATCACTG 110
QY 383 GCAACATCTACATCTACATGACAGTACTGCGGGGACCCAGCGGTCTGAGACCTCC 442
Db 109 GCAACATCTACATCTACATGACAGTACTGCGGGGACCCAGCGGTCTGAGACCTCC 50
QY 443 CAGTACCCCGCAACCTCCATACCCCAT 470
Db 49 CAGTACCCCGCAACCTCCATTCCTCCCT 22

RESULT 10
US-09-907-372-16
; Sequence 16, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.

; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 702245091H1
US-09-907-372-16

Query Match 41.8%; Score 240.4; DB 10; Length 371;
Best Local Similarity 82.3%; Pred. No. 1.7e-64;
Matches 288; Conservative 0; Mismatches 61; Indels 1; Gaps 1;
QY 227 GGCTCCCGCAGCCCGAGTTTGGAGGAGGGTGGCGCAACAGCAGAGTCTCTGGACC 286
Db 4 GACTCCCGCAGCAGCCCGCAAGTCTGGAGGAAGAGATGCTACAGCAGCAGAGTCTCTCAGCC 63
QY 287 TGACACGAGGAGCCGAGTGGAAACCGGGGAGCAGAGCCAGGTGGCCCGCAGGTACCAATG 346
Db 64 AGGCCAGAGAGCTGGAGCCTGAGCTCCAGAAAGCCAGGTGGCCCGCAGGTACCAATG 123
QY 347 GCATTTCATGTCACCGGGGTCTATGACTACTGATCTGCAACATCTACATCTACAATGAC 406
Db 124 GCATTTCATGTCACCGGGGTCTGACTGTGACTGTGCAACATCTACATCTACAATGAC 183
QY 407 CAGTACTGGGGGAGCAGCCCGGTCTGGAGACCTCCCGAGTACCCCGAACCCTCCATACC 466
Db 184 CAGTACTGGGGGAGCAGCAGGAGCCCTGGAGACCCCTGCTTCCCGAGAGCCTCCATACC 243
QY 467 CCATTCCGGAAGAGGGGAGCCCTGGCGCTCCCGGGTCTCTACACCCCGCAGGAAGATG 526
Db 244 CCACCCCTGAAGAGGTGCGCCCTGGCCCTCTCTGAGTCTCTACACCCCTACAGGAGATG 303
QY 527 GCAAGCTTGGCACCTAGCGGAGACAGAG-CACTGTGTGGTGGCACACCCCTC 575
Db 304 GCAAGCTTGGCACCTAGCGGAGACAGAGCCCTGGGGTGGCCATGCCCTC 353

RESULT 11
US-09-907-372-10
; Sequence 10, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 697459H1
; NAME/KEY: unsure
; LOCATION: 76, 131
; OTHER INFORMATION: a, t, c, g, or other
US-09-907-372-10

Query Match 34.5%; Score 198.4; DB 10; Length 219;
Best Local Similarity 98.5%; Pred. No. 1.3e-51;
Matches 199; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 374 CTATCAGTGGCAACATCTACATCTACATGACAGTACTGGGGGAGACACCGGGTCTTG 433
Db 9 CGATCAGTGGCAACATCTACATCTACATGACAGTACTGGGGGAGACACCGGGTCTTG 68
QY 434 GAGACCTCCAGCTACCCGGAACCTTCATACCCATTCCTGAGAGGGGAGCCCTGGCC 493
Db 69 GAGACCTCCAGCTACCCGGAACCTTCATACCCATTCCTGAGAGGGGAGCCCTGGCC 128
QY 494 CTCGCGGGCTCTACACCCGGAACCTTCATACCCATTCCTGAGAGGGGAGCCCTGGCC 553
Db 129 CTNCGGGGCTCTACACCCGGAACCTTCATACCCATTCCTGAGAGGGGAGCCCTGGCC 188
QY 554 AGCACTGTGGTGGCCACACCTTC 575
Db 189 AGCACTGTGGTGGCCACACCTTC 210

RESULT 12
US-09-907-372-11
; Sequence 11, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 11
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 3321983H1
US-09-907-372-11

Query Match 29.9%; Score 172; DB 10; Length 279;
Best Local Similarity 100.0%; Pred. No. 1.8e-43;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 GACCAGTACTGGGGGAGACCGGCTCTGAGAGCTCCAGCTACCCGGAACCTCCAT 463
Db 1 GACCAGTACTGGGGGAGACCGGCTCTGAGAGCTCCAGCTACCCGGAACCTCCAT 60
QY 464 ACCCATTCGCGAAGAGGGGAGCCCTGGCCCTCCCGGGCTCTACACCCGAGGAAG 523
Db 61 ACCCATTCGCGAAGAGGGGAGCCCTGGCCCTCCCGGGCTCTACACCCGAGGAAG 120
QY 524 ATGCAAGGCTTGCCACTAGCGGAGACAGCACTGTGGTGGCCACACCTC 575
Db 121 ATGCAAGGCTTGCCACTAGCGGAGACAGCACTGTGGTGGCCACACCTC 172

RESULT 13
US-09-907-372-13
; Sequence 13, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 206
; TYPE: DNA
; ORGANISM: Rattus norvegicus
```

```
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 700302531H1
US-09-907-372-13

Query Match 25.1%; Score 144.2; DB 10; Length 206;
Best Local Similarity 81.5%; Pred. No. 6e-35;
Matches 167; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 280 CTGAGCCTGACACGAGGAGCCGAGTTGGANCCCGGGGAGAGAGCAGGTGGCCACGGT 339
Db 2 CTGATCCAGCGGAGGAGCTGGAGGCTGAGCCTGGGGAACATGCGCAGGTGGCCACGGT 61
QY 340 ACCAATGGCATTTCACTCACCGCGGGTCTATGACTATCTATGCAACATCTACATCTAC 399
Db 62 GCGAATGGCATTTCACTGACCGGAGGCTCTGTGACTGTACCGGCAATATCTACATATAC 121
QY 400 AATGGACAGTACTGGGGGAGACCGGGTCTCTGAGAGCTCCAGCTACCCGGAACCT 459
Db 122 AATGGGCAAGTGTGGGGGGAACACGGGCGCTCTGAGAGACCTCCAGCTCCCGCTGAGCCT 181
QY 460 CCATACCCCATTCGCCGAAGAGGGGG 484
Db 182 CCATACCCCATTCGCCGAAGAGGGAG 206

RESULT 14
US-09-907-372-12/c
; Sequence 12, Application US/09907372
; Patent No. US20020068242A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti G.
; APPLICANT: Warren, Bridget A.
; TITLE OF INVENTION: TNF RECEPTOR 2 RELATED PROTEIN VARIANT
; FILE REFERENCE: PC-0050 US
; CURRENT APPLICATION NUMBER: US/09/907,372
; CURRENT FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 862
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020068242A1 8576918T1
US-09-907-372-12

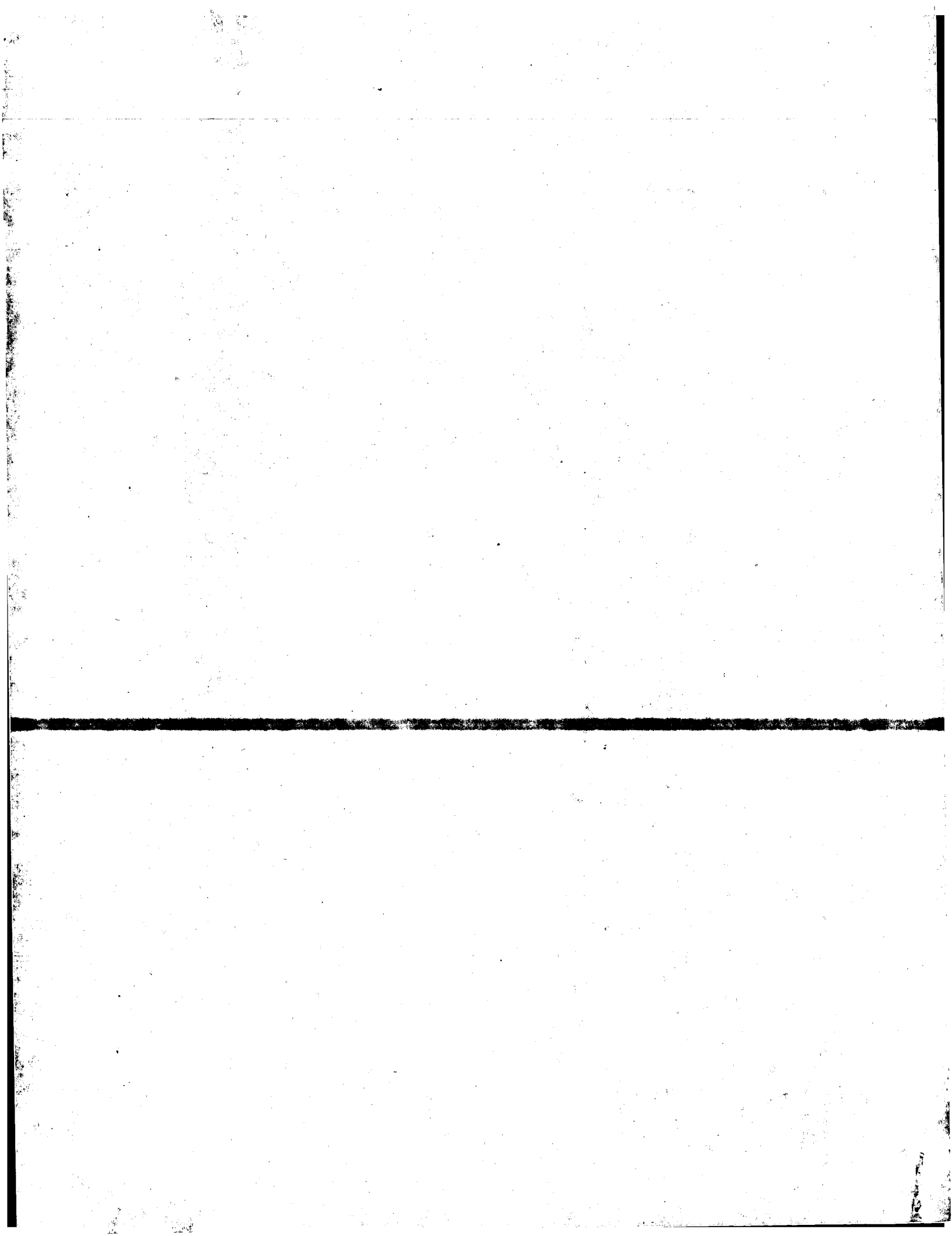
Query Match 17.7%; Score 101.8; DB 10; Length 862;
Best Local Similarity 77.8%; Pred. No. 9.1e-22;
Matches 200; Conservative 0; Mismatches 47; Indels 10; Gaps 6;

QY 328 GTGGCCGACGGTACCAATGGCATTCATGTCACCGCGGGTCTATGACTATCATCTGCAAC 387
Db 853 GTGGCCGCGGTACCCAAAGGCAATAAGGT-ACCGCGGGTTCAAAGAACATCAGCGGCAA 795
QY 388 A-TCTACATCTACAAATGGACCA--GTACTGGGGGAGACACCGGGTCTCTGAGACCTCC 443
Db 794 AGTCGAATTCACAAAGGAGACCAAGTACTGGGGGAGACACCGGGTCTCTGAGACCTCC 735
QY 444 AGCTACCCCGGAACC--TCCATATCCCATTCCTCCGAGAGGGGAGCCCTGGCCCTCCCGG 501
Db 734 AGGCTACCCCGGAGCCCTTCCATACCCCATTCCTCCGAGAGGGGAGCCCTGGCCCTCCCGG 675
QY 502 CTCTCTACACCCGAGGAGGAGTGG--CAAGGCTTGGCACTAGCGGAGACAGAGC-AC 558
Db 674 CTCTCTACACCCGAGGAGGAGTGGCAAGGCTTTGGCACTAGCGGAGACAGAGCAAC 615
QY 559 TGTGTTGCCACACCCCTC 575
Db 614 TGTGTTGCCACACCCCTC 598
```

RESULT 15
US-09-867-701-4881
; Sequence 4881, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4881
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-4881

Query Match 10.6%; Score 61.2; DB 10; Length 481;
Best Local Similarity 95.5%; Pred. No. 2.5e-09;
Matches 63; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CTGGTGGAGGCGACTCCAGGCGACTGCCAGTCCGACACACAACTGCAAAAATCCATTAGAG 60
|||||
Db 416 CTGGTGGAGGCGAGATCCAGGCGACTGCACAGTCCACACAACTGCAAAAATCCATTAGAG 475
|||||
QY 61 CCACTG 66
|||||
Db 476 CCACTG 481

Search completed: April 15, 2003, 18:45:50
Job time : 47.0647 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2003, 10:39:00 : Search time 5.05967 Seconds
(without alignments)
526.717 Million cell updates/sec

Title: US-09-917-372-1_COPY_216_235
Perfect score: 108
Sequence: 1 PLPPMSGSLKRRPQGECP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	80	74.1	305	ABP41926	Human ovarian anti
2	52	48.1	570	ABG26356	Novel human diagno
3	51	47.2	127	AAW933585	Human polypeptide,
4	51	47.2	208	ABG12712	Novel human diagno
5	51	47.2	388	ABG70955	S. spinosa protein
6	49	45.4	80	AAU47321	Propionibacterium
7	49	45.4	207	AAU68531	Human novel cytoki
8	49	45.4	601	AAW43399	Human cancer assoc
9	49	45.4	668	ABW59618	Drosophila melanog
10	49	45.4	891	ABW62720	Drosophila melanog

11	49	45.4	2153	22	AAU33195	Novel human secret
12	48.5	44.9	304	22	ABG07179	Novel human diagno
13	48	44.4	108	21	ABW41588	Human OREF1352
14	47	43.5	159	21	AGC22742	Zea mays protein f
15	47	43.5	1542	22	ABW71456	Drosophila melanog
16	46.5	43.1	206	22	ABG13470	Novel human diagno
17	46	42.6	77	22	AAW85724	Human immune/haema
18	46	42.6	197	18	AAW23220	Extracellular doma
19	46	42.6	197	20	AAV31326	Human lymphotoxin
20	46	42.6	205	22	ABW17052	Human nervous syst
21	46	42.6	235	22	AAW25470	Human protein sequ
22	46	42.6	371	21	AAW57103	Human prostate can
23	46	42.6	425	22	AAW75454	Human colon cancer
24	46	42.6	471	21	AAW31657	Arabidopsis thalia
25	46	42.6	788	22	AAW39095	Human polypeptide
26	46	42.6	788	23	AAW23718	Human GNK interact
27	46	42.6	820	22	AAW40881	Human polypeptide
28	46	42.6	1064	17	AAW96037	Protein tyrosine k
29	46	42.6	1124	23	AAW96931	Human Jak3 protein
30	46	42.6	1124	23	AAW96933	Human Jak3 protein
31	46	42.6	1124	23	AAW96934	Human Jak3 protein
32	46	42.6	1124	23	AAW96935	Human Jak3 protein
33	46	42.6	1124	23	AAW96936	Human Jak3 protein
34	46	42.6	1124	23	AAW96937	Human Jak3 protein
35	46	42.6	1124	23	AAW96938	Human Jak3 protein
36	46	42.6	1124	23	AAW96939	Human Jak3 protein
37	46	42.6	1124	23	AAW96940	Human Jak3 protein
38	46	42.6	1124	23	AAW96941	Human Jak3 protein
39	46	42.6	1124	23	AAW96942	Human Jak3 protein
40	46	42.6	1124	23	AAW96943	Human Jak3 protein
41	46	42.6	1124	23	AAW96944	Human Jak3 protein
42	46	42.6	1124	23	AAW96945	Human Jak3 protein
43	46	42.6	1124	23	AAW96946	Human Jak3 protein
44	46	42.6	1129	22	ABW20477	Novel human diagno
45	46	42.6	1461	19	AAW64468	Human secreted pro

ALIGNMENTS

RESULT 1

ABP41926
ID ABP41926 standard; Protein; 305 AA.
XX
AC ABP41926;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HSABJ44, SEQ ID NO:3058.

Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;
Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;
Ovarian cancer; breast cancer; tumour; reproductive system disorder;
Infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
Inflammatory condition; immune disorder; blood disorder;
cardiovascular disorder; respiratory disorder; neurological disorder;
gastrointestinal disorder; urinary system disorder; drug screening;
gene therapy; chromosome mapping; forensic analysis;
antibody preparation; cytostatic; immunomodulatory; neuroprotective;
antiinflammatory; gynaecological; reproductive; chromosome 12p13.

OS Homo sapiens.

XX
PN WO200200677-A1.

XX
PD 03-JAN-2002.

XX
PF 07-JUN-2001; 2001WO-US18569.

XX
PR 07-JUN-2000; 2000US-209467P.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX

Blrse CE, Rosen CA;
WPI; 2002-147878/19.
N-PSDB; ABQ55003.

Isolated nucleic acid molecules encoding novel ovarian polypeptides, useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and neurological diseases -

Claim 11; SEQ ID No 3058; 2922pp; English.

The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 305 AA;

Query Match 74.1%; Score 80; DB 23; Length 305;
Best Local Similarity 35.7%; Pred. No. 0.0012;
Matches 20; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 PLPPMS-----GSLKKRRPOGEP 20
|||||
DB 86 PLPPMSGTMLMLVLLPFLAFLLATVFSCIWKSHPSLCRKLGSLKKRRPOGEP 141
|||||

RESULT 2
ABG26356
ID ABG26356 standard; Protein; 570 AA.
XX AC
XX ABG26356;
XX DT
XX DE
XX DE
XX DE
XX KW
XX KW
XX KW
XX OS
XX XX
XX PN
XX PD
XX XX
XX PF

Novel human diagnostic protein #26347.
Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
Homo sapiens.
WO200175067-A2.
11-OCT-2001.
30-MAR-2001: 2001WO-US08631.


```

PR 02-MAY-2000; 2000JP-0183765.
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
DR N-PSDB; AAK94517.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
PT
XX Claim 8; SEQ ID NO 3381; 1380pp + sequence listing; English.
XX
XX The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
XX Sequence 127 AA;
SQ
Query Match 47.2%; Score 51; DB 22; Length 127;
Best Local Similarity 50.0%; Pred. No. 7.9;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
QY 1 PLPPEMSGSLKRRPQCEGP 20
DB 61 PLPPEVRGSLPEGAPWSRAP 80
RESULT 4
ABG12712
ID ABG12712 standard; Protein; 208 AA.
AC
AC ABG12712;
XX
XX 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #12703.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX .30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX N-PSDB; AAS76899.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
PT
PR 02-MAY-2000; 2000JP-0183765.
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
DR N-PSDB; AAK94517.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
PT
XX Claim 8; SEQ ID NO 3381; 1380pp + sequence listing; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC the polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 208 AA;
SQ
Query Match 47.2%; Score 51; DB 22; Length 208;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 1 PLPPEMSGSLKRRPQCE 18
DB 179 PLPSAMSGSLRPLPEAE 196
RESULT 5
AAB70955
ID AAB70955 standard; Protein; 388 AA.
AC
AC AAB70955;
XX
XX 28-AUG-2001 (first entry)
XX
XX S. spinosa protein fragment encoded by ORF8, SEQ ID 22.
XX
XX Forosamine; trimethylrhamnose; polyketide synthase; biosynthesis;
XX spinosyn; polyketide aglycone; transgenic plant; insect resistance;
XX macrolide; insecticidal; C-C cyclising enzyme.
XX
XX Saccharopolyspora spinosa.
XX
XX DE19957268-A1.
XX
XX 08-MAR-2001.
XX
XX 29-NOV-1999; 99DE-1057268.
XX
XX 27-AUG-1999; 99DE-1040596.
XX
XX (FARB ) BAYER AG.
XX
XX Eberz G, Moehrle V, Froede R, Velten R, Salas JA;
XX
XX WPI; 2001-267102/28.
XX
XX N-PSDB; AAF88325.
XX
XX New nucleic acid encoding enzymes for spinosyn biosynthesis, useful for
XX recombinant production of insecticidal spinosyns and their derivatives
XX
XX Claim 39; Page 131-132; 354pp; German.
PS

```

XX This invention describes a novel method nucleic acid (I) and its encoded
 CC polypeptide (II) containing at least one region that encodes an enzymatic
 CC activity involved in biosynthesis of spinosyns. (I) are used (i) to
 CC identify, inactivate or modulate genes involved in the biosynthesis of
 CC (II); (ii) to generate a library of polypeptide synthases; (iii) for
 CC adding forosamine or trimethylrhamsone to a spinosyn or polypeptide
 CC aglycone; and (iv) for recombinant production of the corresponding
 CC derivatives, including production of (II), their precursors or
 CC and thus have increased resistance to insects. (I) are also useful as
 CC markers for sequencing of the *Saccharopolyspora spinosa* genome. (II) are
 CC macrolides with insecticidal, but not antibacterial, activity, and can
 CC also be used to raise specific antibodies, useful for identifying
 CC expression clones in a gene bank. Cells transformed with (I) may produce
 CC (II) at significantly increased levels or produce new derivatives of
 CC (II). This sequence represents an *S. spinosa* C-C cyclising enzyme.
 XX

SQ Sequence 388 AA;

Query Match 47.2%; Score 51; DB 22; Length 388;
 Best Local Similarity 60.0%; Pred. No. 25;
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRP 15
 ||||:|:|:|
 Db 27 PLPPGVPTLLRAP 41

RESULT 6

AAU47321
 ID AAU47321 standard; Protein; 80 AA.

AC AAU47321;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #8217.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypervitosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59538.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris -

XX Example 1; SEQ ID No 8516; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 80 AA;

Query Match 45.4%; Score 49; DB 22; Length 80;
 Best Local Similarity 50.0%; Pred. No. 9.7;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 PLPPEMSGSLKRRPQGE 18
 ||||:|:|:|
 Db 52 PSPPELAGARUGRHYQQG 69

RESULT 7

AAU68531

ID AAU68531 standard; Protein; 207 AA.

AC AAU68531;

DT 16-JAN-2002 (first entry)

DE Human novel cytokine encoded by cDNA 790CIP2C_2 #1.

XX Human; cytokine; cell proliferation; cell differentiation;
 KW antiinflammatory; stem cell growth factor; activin; inhibin; cancer;
 KW nervous system disease; neuropathy; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; spinal cord disorder;
 KW head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
 KW platelet disorder; thrombocytopaenia; stem cell disorder;
 KW aplastic anaemia; tissue regeneration; wound healing; ulcer;
 KW osteoporosis; osteoarthritis; bone degenerative disorder;
 KW periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
 KW severe combined immunodeficiency; infection; autoimmune disorder;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
 KW asthma; coagulation disorder; haemophilia; sepsis; nephritis;
 KW inflammatory bowel disease; food supplement; immunogen.

OS Homo sapiens.

XX WO200175093-A1.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US10484.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX 22-SEP-2000; 2000US-0668680.

XX 23-OCT-2000; 2000US-0695618.

XX 30-NOV-2000; 2000US-0728711.

XX 14-MAR-2001; 2000US-0728711.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;

PI Yang Y, Zaho QA, Chen R, Wang D, Goodrich RW, Liu C, Drmanac RT;

DR WPI; 2001-G26432/72.
 XX N-PSDB; AA559823.
 XX
 PT New polypeptides and nucleic acids, useful for diagnosis, treatment of
 PT inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone
 PT degenerative disorders, cancer and promoting wound healing
 XX
 XX
 PS Claim 20; Page 248; 336pp; English.
 XX
 XX The invention relates to isolated human polypeptides (which may be
 CC cytokines) and the polynucleotides encoding them. The protein is useful
 CC for identifying a compound which binds to it (e.g. modulators, agonists
 CC and antagonists). The polynucleotides are useful as an array for mismatch
 CC detection. The proteins and nucleic acids are useful as nutritional
 CC sources or supplements. The protein exhibits activity relating
 CC to cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity, immune stimulating or immune
 CC suppressing and activin or inhibin related activities. The proteins (and
 CC antibodies raised against them) and nucleic acids are therefore useful in
 CC the diagnosis and treatment of diseases and disorders such as cancer,
 CC Alzheimer's, Parkinson's disease, Huntington's disease, ankyrotrophic
 CC lateral sclerosis, spinal cord disorders, head trauma, cerebrovascular
 CC diseases, stroke, myeloid or lymphoid cell disorders, platelet disorders,
 CC thrombocytopaenia, stem cell disorders, aplastic anaemia, for
 CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
 CC growth, and in tissue repair, healing of burns, incisions, ulcers, or
 CC treating osteoporosis, osteoarthritis, bone degenerative disorders, or
 CC periodontal disease, lung or liver fibrosis, reperfusion injury in
 CC various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis), allergic reactions and conditions,
 CC such as asthma or other respiratory problems, coagulation disorders,
 CC haemophilia, septic shock, sepsis, arthritis, nephritis and inflammatory
 CC bowel disease, viral infection and are useful in altering bodily
 CC characteristics. The present sequence represents a novel protein of the
 CC invention.
 XX
 XX Sequence 207 AA;
 SQ
 Query Match 45.4%; Score 49; DB 22; Length 207;
 Best Local Similarity 31.2%; Pred. No. 26;
 Matches 10; Conservative 5; Mismatches 5; Indels 12; Gaps 1;
 QY 1 PLPPEMSGSL-----LKRPPQGEPP 20
 ||||: :| :|:| :|:|
 Db 168 PLPDTAGRAQHSQSPRRALPAKRPQGP 199
 RESULT 8
 AAB43399
 ID AAB43399 standard; Protein; 601 AA.
 XX
 XX AAB43399;
 AC
 XX
 XX 08-FEB-2001 (first entry)
 DT
 XX Human cancer associated protein sequence SEQ ID NO:844.
 DE
 XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antiidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; organ rejection;
 KW allergic reaction; graft versus host disease; autoimmune disorder;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 XX Homo sapiens.
 OS
 XX

PN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US05882.
 PF
 XX 12-MAR-1999; 99US-0124270.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2000-587533/55.
 DR N-PSDB; AAC77608.
 XX
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 PT
 XX Claim 11; Page 1396-1398; 2352pp; English.
 PS
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antiidiabetic; antiasthmatic; antirheumatic; antirheumatic; antithyroid;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX
 XX Sequence 601 AA;
 SQ
 Query Match 45.4%; Score 49; DB 21; Length 601;
 Best Local Similarity 47.4%; Pred. No. 76;
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 2 LPPEMSGSLKRRPQGEPP 20
 ||||: :| :|:| :|:|
 Db 194 LPPPHSSGFLGSKPGEPP 212
 RESULT 9
 ABB59618
 ID ABB59618 standard; Protein; 668 AA.
 XX
 XX ABB59618;
 AC
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 5646.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW
 KW Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 PF

CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

XX Sequence 2153 AA;

Query Match 45.4%; Score 49; DB 22; Length 2153;
 Best Local Similarity 47.4%; Pred. No. 2.8e+02;
 Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 LPPEMSGSLKRRPQEGP 20
 ||| | | :|||
 Db 1562 LPPHSSGFLGSKRPGPG 1580

RESULT 12

ABG07179
 ID ABG07179 standard; Protein: 304 AA.

XX AC ABG07179;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #7170.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX PF WPI; 2001-639362/73.

XX PR N-PSDB; AAS71366.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX PS Claim 20; SEQ ID No 37538; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 304 AA;

Query Match 44.9%; Score 48.5; DB 22; Length 304;
 Best Local Similarity 52.0%; Pred. No. 45;
 Matches 13; Conservative 1; Mismatches 4; Indels 7; Gaps 2;

QY 1 PLPPEM-SGSLL-----RRRQGE 18
 ||||| | | :|||

Db 250 PLPPEFSPGALSSCRXGNRRPEGE 274

RESULT 13

AAB41588

ID AAB41588 standard; Protein: 108 AA.

XX AC AAB41588;

XX DT 08-FEB-2001 (first entry)

XX DE Human ORFX ORF1352 polypeptide sequence SEQ ID NO:2704.

XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX OS Homo sapiens.

XX PN WO200058473-A2.

XX PD 05-OCT-2000.

XX PF 31-MAR-2000; 2000WO-US08621.

XX PR 31-MAR-1999; 99US-0127607.

XX PR 02-APR-1999; 99US-0127636.

XX PR 05-APR-1999; 99US-0127728.

XX PR 30-MAR-2000; 2000US-0540763.

XX PA (CURA-) CURAGEN CORP.

XX PI Shinkets RA, Leach M;

XX PF WPI; 2000-602362/57.

XX PR N-PSDB; AAC75797.

XX PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -

XX PS Claim 11; Page 1941; 5507pp; English.

XX CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;

CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

SQ Sequence 108 AA;

Query Match 44.4%; Score 48; DB 21; Length 108;
 Best Local Similarity 63.2%; Pred. No. 18;
 Matches 12; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

QY 2 LPPEMSGILKRRPOGEP 20
 ||||| ||||| |||||
 Db 66 LPPE--GILLPLLLGEP 82

RESULT 14

AAG22742
 ID AAG22742 standard; Protein; 159 AA.

XX AAG22742;

AC AAG22742;

XX 17-OCT-2000 (first entry)

XX Zea mays protein fragment SEQ ID NO: 25788.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.

XX Zea mays subsp. mays.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132487.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 18-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 22-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 23-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
02-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
03-AUG-1999; 99US-0147038.
04-AUG-1999; 99US-0147204.
05-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147935.
09-AUG-1999; 99US-0147935.
10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-0149175.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149902.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
26-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
27-AUG-1999; 99US-0151303.
30-AUG-1999; 99US-0151438.
31-SEP-1999; 99US-0151930.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.
15-SEP-1999; 99US-0154018.
16-SEP-1999; 99US-0154039.
20-SEP-1999; 99US-0154779.
22-SEP-1999; 99US-0155139.
23-SEP-1999; 99US-0155486.
24-SEP-1999; 99US-0155659.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.
05-OCT-1999; 99US-0157753.
08-OCT-1999; 99US-0157865.
07-OCT-1999; 99US-0158029.
08-OCT-1999; 99US-0158232.
12-OCT-1999; 99US-0158369.
13-OCT-1999; 99US-0159293.
13-OCT-1999; 99US-0159294.
13-OCT-1999; 99US-0159295.
14-OCT-1999; 99US-0159329.
14-OCT-1999; 99US-0159330.
14-OCT-1999; 99US-0159331.
14-OCT-1999; 99US-0159637.
18-OCT-1999; 99US-0159638.
18-OCT-1999; 99US-0159584.
21-OCT-1999; 99US-0160741.
21-OCT-1999; 99US-0160767.
21-OCT-1999; 99US-0160768.
21-OCT-1999; 99US-0160770.
21-OCT-1999; 99US-0160814.
21-OCT-1999; 99US-0160815.
22-OCT-1999; 99US-0160980.
22-OCT-1999; 99US-0160981.

22-OCT-1999; 99US-0160989.
25-OCT-1999; 99US-0161404.
25-OCT-1999; 99US-0161405.
25-OCT-1999; 99US-0161406.
26-OCT-1999; 99US-0161359.
26-OCT-1999; 99US-0161360.
26-OCT-1999; 99US-0161361.
28-OCT-1999; 99US-0161920.
28-OCT-1999; 99US-0161992.
28-OCT-1999; 99US-0161993.
29-OCT-1999; 99US-0162142.

Query Match 43.5%; Score 47; DB 21; Length 159;
Best Local Similarity 47.4%; Pred. No. 38;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 PLPPMSGSLKKRRPQEG 19
| | | | | | | | | |
Db 16 PPPQSSRRSRNPRGTG 34

RESULT 15

ABB71456

ID ABB71456 standard; Protein; 1542 AA.

XX AC ABB71456;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 41160.

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

XX WC200171042-A2.

PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL15559.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX Disclosure; SEQ ID NO 41160; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1542 AA;

Query Match 43.5%; Score 47; DB 22; Length 1542;

Best Local Similarity 52.9%; Pred. No. 3.9e+02;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy 4 PMSGSLKKRRPQGE 20
| | : | | | | |
Db 1016 PYMPDHMYPRRQAGP 1032

Search completed: April 8, 2003, 10:49:07
Job time : 7.05967 secs